

76/14

From: Davis, Natalie  
Sent: Thursday, September 19, 2002 3:18 PM  
To: STIC-Biotech/ChemLib  
Subject: 09/418887

Please search SEQ ID NO: 2 for 09/807470.

Thanks,

Natalie A. Davis, PhD  
Patent Examiner  
Art Unit 1642  
CM1, Rm 8B13  
Mailbox 8E12  
Ph (703) 308-6410

Edward Hart  
Technical Info. Specialist  
STIC/Biotech  
CMI 6B02 Tel: 305-9203

Pending Nucleic Acid and/or Pending Amino Acid database searches now generate two sets of results. These databases were split into two parts to reduce the time needed to update the databases daily. The split freed up more machine time for processing searches.

Searches run against the Nucleic Acid Pending database produce two sets of results, with the extensions, .rnpm and .rnpn

Searches run against the Amino Acid Pending database produce two sets of results, with the extensions, .rapm and .rapn

***The Pending database search results should not be left in the case because they contain data that is confidential.***

## TYPE OF SEARCH:

Searcher: \_\_\_\_\_  
Phone: \_\_\_\_\_  
Location: \_\_\_\_\_  
Date Picked Up: 9/20/02  
Date Completed: 9/20/02  
Searcher Prep/Review: \_\_\_\_\_  
Clerical: \_\_\_\_\_  
Online time: \_\_\_\_\_

NA Sequences: \_\_\_\_\_  
AA Sequences: 1  
Structures: \_\_\_\_\_  
Bibliographic: \_\_\_\_\_  
Litigation: \_\_\_\_\_  
Full text: \_\_\_\_\_  
Patent Family: \_\_\_\_\_  
Other: \_\_\_\_\_

## VENDOR/COST (where applic.)

STN: \_\_\_\_\_  
DIALOG: \_\_\_\_\_  
Questel/Orbit: \_\_\_\_\_  
DRLink: \_\_\_\_\_  
Lexis/Nexis: \_\_\_\_\_  
Sequence Sys.: Q2  
WWW/Internet: \_\_\_\_\_  
Other (specify): \_\_\_\_\_

GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 20, 2002, 14:34:08 ; Search time 32.48 seconds  
(without alignments)  
1241.373 Million cell updates/sec

Title: US-09-807-470-2

Perfect score: 1892

Sequence: 1 MGLRKKRNARNPVLSHEFV.....NGVENPNRIDSPPKKKEKAP 363

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 747574 seqs, 11107396 residues

Total number of hits satisfying chosen parameters: 747574

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

- A.Geneseq\_032802:\*
- 1: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1980.DAT:\*
  - 2: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1981.DAT:\*
  - 3: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1982.DAT:\*
  - 4: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1983.DAT:\*
  - 5: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1984.DAT:\*
  - 6: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1985.DAT:\*
  - 7: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1986.DAT:\*
  - 8: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1987.DAT:\*
  - 9: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1988.DAT:\*
  - 10: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1989.DAT:\*
  - 11: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1990.DAT:\*
  - 12: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1991.DAT:\*
  - 13: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1992.DAT:\*
  - 14: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1993.DAT:\*
  - 15: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1994.DAT:\*
  - 16: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1995.DAT:\*
  - 17: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1996.DAT:\*
  - 18: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1997.DAT:\*
  - 19: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1998.DAT:\*
  - 20: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1999.DAT:\*
  - 21: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA2000.DAT:\*
  - 22: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA2001.DAT:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Match	Length	ID	Description
1	1892	100.0	363	21	Rat WAR-1 amino ac
2	1892	100.0	363	22	Rat WAR-1 protein
3	1463	77.3	369	21	Human WAR-1 amino
4	1463	77.3	369	22	Novel human diagno
5	1463	77.3	369	22	Human polypeptide,
6	1462	77.3	369	22	Human cancer assoc
7	1134.5	60.0	416	21	Human prostate can
8	922	48.7	304	20	Human reproductive
9	739	39.1	384	22	Drosophila melanog
10	566	29.9	368	22	ABB65847
11	566	29.9	368	22	ABB66114

12	566	29.9	1575	22	ABB67362	Drosophila melanog
13	414.5	21.9	125	21	AAG00189	Human secreted pro
14	148.5	7.8	411	17	AA86810	Saccharomyces cere
15	138.5	7.3	394	20	AA100876	Human LAPH-1 prote
16	138.5	7.3	394	22	AAW78909	Human protein seq
17	138.5	7.3	394	22	AA893884	Human secreted pro
18	133.5	7.1	191	21	AA824489	Human polypeptide
19	117	6.2	488	22	AAW42028	Human protein seq
20	117	6.2	536	22	AAW79893	Arabidopsis thalia
21	115.5	6.1	296	21	AAG44196	Arabidopsis thalia
22	114.5	6.1	280	21	AAG44197	Novel signal trans
23	111.5	5.9	158	22	AAU17576	Novel signal trans
24	111.5	5.9	181	22	AAU17201	Human olfactory re
25	111	5.9	312	22	AAG72224	Human OR-like poly
26	111	5.9	312	22	AAG72549	Human LAPH-2 prote
27	110	5.8	380	20	AAV00877	Human hydrophobic
28	110	5.8	380	22	AA888560	Human apoptosis pr
29	108	5.7	311	22	AAU00782	Human cancer assoc
30	106.5	5.6	266	21	AA843806	Arabidopsis thalia
31	105.5	5.6	310	21	AAG41779	Human secreted pro
32	101.5	5.4	150	21	AA834938	Drosophila melanog
33	98	5.2	232	22	AB867693	Zea mays protein f
34	97.5	5.2	313	21	AA833404	Chlamydia trachoma
35	97.5	5.2	397	20	AAV37316	Human polypeptide
36	97.5	5.2	533	22	AAW40417	Novel human diagno
37	96.5	5.1	431	22	ABG30056	Arabidopsis thalia
38	96	5.1	254	21	AA841780	Human G protein-co
39	93	4.9	313	21	AAV90877	Human olfactory re
40	93	4.9	313	22	AAU24525	Human olfactory re
41	92	4.9	313	22	AAG72073	Human olfactory re
42	91	4.8	307	22	AAG72203	Human olfactory re
43	91	4.8	315	22	AAG72870	Human olfactory re
44	91	4.8	324	22	AAW71539	Human olfactory re
45	90	4.8	309	22	AAU10312	G-protein coupled

## ALIGNMENTS

RESULT 1  
AAV98146  
ID AAV98146 standard; Protein; 363 AA.  
XX  
AC AAV98146;  
XX  
DT 22-AUG-2000 (first entry)  
XX  
DE Rat WAR-1 amino acid sequence.  
XX  
KW Endoplasmic reticulum; WAR-1; cancer cell proliferation inhibitor;  
KW diagnosis; cancer; sarcoma; rat.  
XX  
OS Rattus norvegicus.  
XX  
PN WO200022123-A1.  
XX  
PD 20-APR-2000.  
XX  
PF 13-OCT-1999; 99WO-JP05631.  
XX  
PR 13-OCT-1998; 98JP-0290711.  
XX  
(SUMU ) SUMITOMO PHARM CO LTD.  
XX  
PI Tohdoh N, Yoshima T, Komiya K, Tojo S, Nemoto K, Ishikawa H;  
PI Okuyama H;  
XX  
DR WPI; 2000-317980/27.  
DR N-PSDB; AAA38012.  
XX  
PT Endoplasmic reticulum protein WAR-1 which inhibits cancer cell  
PT proliferation for use in treatment and diagnosis of cancer including  
PT sarcomas of high malignancy

XX  
PS Claim 1; Fig 2; 89pp; Japanese.  
XX  
CC This sequence represents an endoplasmic reticulum protein (WAR-1) amino  
CC acid sequence. The invention includes rat and human WAR-1 sequences,  
CC expression vectors containing the DNA, cells transformed with the  
CC expression vector, antibodies against WAR-1, and probes and primers which  
CC hybridize to the DNA encoding WAR-1. The WAR-1 protein inhibits the  
CC proliferation of cancer cells, and is used in the treatment and diagnosis  
CC of cancers including highly malignant sarcomas.  
XX  
SQ Sequence 363 AA;

Query Match 100.0%; Score 1892; DB 21; Length 363;  
Best Local Similarity 100.0%; Pred. No. 3.8e-213;  
Matches 363; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 MGLRKNARNPPVLSHEPFVQNHADVMSCVGMFFVLGLMFEFTAEMSIVFLTLQHGVPVP 60  
DB 1 MGLRKNARNPPVLSHEPFVQNHADVMSCVGMFFVLGLMFEFTAEMSIVFLTLQHGVPVP 60  
QY 61 AGLPSGSRITLYHYGVKDLATVFFVLMVAIIHATIOEYVLDKLSRLQLTKGQNKLINE 120  
DB 61 AGLPSGSRITLYHYGVKDLATVFFVLMVAIIHATIOEYVLDKLSRLQLTKGQNKLINE 120  
QY 121 AQGLSVFYIVSGIWMIIASENCISDPTLLWKSOPHNMTFQMKFFYISOLAYWFHSFP 180  
DB 121 AQGLSVFYIVSGIWMIIASENCISDPTLLWKSOPHNMTFQMKFFYISOLAYWFHSFP 180  
QY 181 ELYFOKVRKQDIPGOLIIYIGLHPIHGAYLLYNHLGLLLMLHYAVELSSVCSLLYF 240  
DB 181 ELYFOKVRKQDIPGOLIIYIGLHPIHGAYLLYNHLGLLLMLHYAVELSSVCSLLYF 240  
QY 241 GDERYOKGLSLWPVIFISGRVTLIVSVTVGLHAGTNRNGNALSNGVNVLAIAKIAVL 300  
DB 241 GDERYOKGLSLWPVIFISGRVTLIVSVTVGLHAGTNRNGNALSNGVNVLAIAKIAVL 300  
QY 301 SSCSQVITVTLTWTWVQLRWLEDANLHVCGRKRSRSGTENGVENPNRIDSPPKKKE 360  
DB 301 SSCSQVITVTLTWTWVQLRWLEDANLHVCGRKRSRSGTENGVENPNRIDSPPKKKE 360  
QY 361 KAP 363  
DB 361 KAP 363

RESULT 2  
AAB70696  
ID AAB70696 standard; Protein; 363 AA.  
XX  
AC AAB70696;  
XX  
DT 17-MAY-2001 (first entry)  
XX  
DE Rat WAR-1 protein sequence SEQ ID NO:2.  
XX  
KW WAR-1; protein screening; endoplasmic reticulum membrane protein;  
KW endoplasmic reticulum membrane transportation; secretory protein;  
KW cell membrane protein; cytosolic; CNS active; antiallergic; cancer;  
KW antirheumatic; nervous system disorder; immune disorder; allergy;  
KW rheumatism; skeletal disorder.  
XX  
OS Rattus sp.  
XX  
PN WO200114582-A1.  
XX  
PD 01-MAR-2001.  
XX  
PF 17-AUG-2000; 2000WO-JP05488.  
XX  
PR 20-AUG-1999; 99JP-0234764.  
XX

PA (SUMU ) SUMITOMO PHARM CO LTD.  
XX  
PI Tohdoh N, Okuyama H, Imamura M, Ishikawa H, Nemoto K;  
XX  
DR WPI; 2001-202940/20.  
XX  
XX N-PSDB; AAF74781.  
PT Transformation of a cell with separate vectors expressing the sense and  
PT antisense strands of WAR-1 DNA for screening secretory and membrane  
PT proteins expressed by the cell -  
XX  
PS Claim 3; Page 60-62; 79pp; Japanese.  
XX

XX The present invention describes a screening method for secretory and  
XX membrane proteins consisting of transformation of a cell with separate  
XX expression vectors for the sense and antisense RNA of DNA encoding an  
XX endoplasmic reticulum membrane protein participating in endoplasmic  
XX reticulum transport of proteins. Also described are: (1) secretory and  
XX cell membrane proteins identified by the screening method; (2) drug  
XX compositions containing these proteins; (3) host cells transformed by  
XX the separate expression vectors of the method; and (4) the preparation  
XX of secretory and cell membrane proteins by culture of the transformants.  
XX The method can be used for the identification and preparation of  
XX proteins for use in the treatment and prevention of diseases such as  
XX cancer, disorders of the nervous system, immune disorders (including  
XX allergies and rheumatism) and skeletal disorders. The present sequence  
XX represents a specifically claimed rat WAR-1 protein from the present  
XX invention.

SQ Sequence 363 AA;  
Query Match 100.0%; Score 1892; DB 22; Length 363;  
Best Local Similarity 100.0%; Pred. No. 3.8e-213;  
Matches 363; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 MGLRKNARNPPVLSHEPFVQNHADVMSCVGMFFVLGLMFEFTAEMSIVFLTLQHGVPVP 60  
DB 1 MGLRKNARNPPVLSHEPFVQNHADVMSCVGMFFVLGLMFEFTAEMSIVFLTLQHGVPVP 60  
QY 61 AGLPSGSRITLYHYGVKDLATVFFVLMVAIIHATIOEYVLDKLSRLQLTKGQNKLINE 120  
DB 61 AGLPSGSRITLYHYGVKDLATVFFVLMVAIIHATIOEYVLDKLSRLQLTKGQNKLINE 120  
QY 121 AQGLSVFYIVSGIWMIIASENCISDPTLLWKSOPHNMTFQMKFFYISOLAYWFHSFP 180  
DB 121 AQGLSVFYIVSGIWMIIASENCISDPTLLWKSOPHNMTFQMKFFYISOLAYWFHSFP 180  
QY 181 ELYFOKVRKQDIPGOLIIYIGLHPIHGAYLLYNHLGLLLMLHYAVELSSVCSLLYF 240  
DB 181 ELYFOKVRKQDIPGOLIIYIGLHPIHGAYLLYNHLGLLLMLHYAVELSSVCSLLYF 240  
QY 241 GDERYOKGLSLWPVIFISGRVTLIVSVTVGLHAGTNRNGNALSNGVNVLAIAKIAVL 300  
DB 241 GDERYOKGLSLWPVIFISGRVTLIVSVTVGLHAGTNRNGNALSNGVNVLAIAKIAVL 300  
QY 301 SSCSQVITVTLTWTWVQLRWLEDANLHVCGRKRSRSGTENGVENPNRIDSPPKKKE 360  
DB 301 SSCSQVITVTLTWTWVQLRWLEDANLHVCGRKRSRSGTENGVENPNRIDSPPKKKE 360  
QY 361 KAP 363  
DB 361 KAP 363

RESULT 3  
AAY98147  
ID AAY98147 standard; Protein; 369 AA.  
XX  
AC AAY98147;  
XX  
DT 22-AUG-2000 (first entry)  
XX

DE Human WAR-1 amino acid sequence.  
 XX Endoplasmic reticulum; WAR-1; cancer cell proliferation inhibitor;  
 KW diagnosis; cancer; sarcoma; human.  
 XX  
 OS Homo sapiens.  
 XX WO200022123-A1.  
 PN 20-APR-2000.  
 XX  
 XX 13-OCT-1999; 99WO-JP05631.  
 XX 13-OCT-1998; 98JP-0290711.  
 XX (SUMI) SUMITOMO PHARM CO LTD.  
 PA Tohdoh N, Yoshima T, Komiya K, Tojo S, Nemoto K, Ishikawa H;  
 PI Okuyama H;  
 XX  
 XX WPI; 2000-317980/27.  
 DR N-PSDB; AAA38013.  
 XX  
 XX Endoplasmic reticulum protein WAR-1 which inhibits cancer cell  
 PT proliferation for use in treatment and diagnosis of cancer including  
 PT sarcomas of high malignancy.  
 XX  
 PS Claim 1; Fig 2; 89pp; Japanese.  
 XX  
 CC This sequence represents an endoplasmic reticulum protein (WAR-1) amino  
 CC acid sequence. The invention includes rat and human WAR-1 sequences,  
 CC expression vectors containing the DNA, cells transformed with the  
 CC expression vector, antibodies against WAR-1, and probes and primers which  
 CC hybridize to the DNA encoding WAR-1. The WAR-1 protein inhibits the  
 CC proliferation of cancer cells, and is used in the treatment and diagnosis  
 CC of cancers including highly malignant sarcomas.  
 XX  
 SQ Sequence 369 AA;

Query Match 77.3%; Score 1463; DB 21; Length 369;  
 Best Local Similarity 75.8%; Pred. No. 9.6e-163;  
 Matches 279; Conservative 43; Mismatches 40; Indels 6; Gaps 5;  
 QY 1 MGLRKNARNPPVLSHEFVQNHADMSCVGMFFVLGLMFECTAEMSIVFLTLQHGVPVP 60  
 DB 1 mglrktstnpvlsqefilqhadiavscvgmffilgfvfegtaeasivfltlqhsavp 60  
 QY 61 -ABGLPSGRTLYHYGVKDLATVFFYMLVAIIHATIQEYVLDKLSRRLQTLKGQKNLN 119  
 DB 61 aaeeqatgsksllyygvkdlatvffymvliiathiqeyvldklnkrmqtkakqknfn 120  
 QY 120 EAGOLSVFVIVSGIMGMIILASENCLSDPTLLKSPQHNMTFQMKFFYISQLAYWPHSF 179  
 DB 121 esqgfsvffvfcwlgwffllisencldptllwkarphsmmtfgmkffysqlaywfhaf 180  
 QY 180 PELYFKVRKQDIPGQIYIGLHFLHIGGAYLLYLNHLGLLMLHYAVELLSSVCSLLY 239  
 DB 181 pelyfaktkkdiprqlvyglnhltgagyllynhlglllvhlvhyfvelslmcmglfy 240  
 QY 240 FGDERYQKGLSWPVIFFISGRVLTIVSVVTVGLHAGT-NRNGNALSGNVNVAAKTAV 298  
 DB 241 fsdekyqkqgislwaflligrlvltivsvltvgfhlagsqnrpdaltgynvnlvlaaklav 300  
 QY 299 LSSSCSIQVYITWLTTLTWLQWLEADANLHV-CGRKRBSR-SRKGTEN--GVENPNRIDS 354  
 DB 301 lssscsiqyvtnltlwlqrwedsnlqascmkkrksrskrtengvgvetsnrvc 360  
 QY 355 PPKKEKA 362  
 DB 361 ppkkeks 368

## RESULT 4

ABGI2234  
 ID ABGI2234 standard; Protein; 369 AA.  
 XX  
 AC ABGI2234;  
 XX  
 XX 18-FEB-2002 (first entry)  
 XX  
 XX Novel human diagnostic protein #12225.  
 XX  
 XX Human; chromosome mapping; gene mapping; gene therapy; forensic;  
 KW food supplement; medical imaging; diagnostic; genetic disorder.  
 KW  
 XX Homo sapiens.  
 OS  
 XX WO200175067-A2.  
 PN 11-OCT-2001.  
 XX  
 XX 30-MAR-2001; 2001WO-US08631.  
 PF  
 XX 31-MAR-2000; 2000US-0540217.  
 PR 23-AUG-2000; 2000US-0649167.  
 XX  
 XX (HYSE-) HYSEQ INC.  
 PA  
 XX Drmanac RT, Liu C, Tang YT;  
 PI  
 XX WPI; 2001-639362/73.  
 DR N-PSDB; AAS76421.  
 XX  
 XX New isolated polynucleotide and encoded polypeptides, useful in  
 PT diagnostics, forensics, gene mapping, identification of mutations  
 PT responsible for genetic disorders or other traits and to assess  
 PT biodiversity.  
 XX  
 PS Claim 20; SEQ ID No 42593; 103pp; English.  
 XX  
 CC The invention relates to isolated polynucleotide (I) and  
 CC polypeptide (II) sequences. (I) is useful as hybridisation probes,  
 CC polymerase chain reaction (PCR) primers, oligomers, and for chromosome  
 CC and gene mapping, and in recombinant production of (II). The  
 CC polynucleotides are also used in diagnostics as expressed sequence tags  
 CC for identifying expressed genes. (I) is useful in gene therapy techniques  
 CC to restore normal activity of (II) or to treat disease states involving  
 CC (II). (II) is useful for generating antibodies against it, detecting or  
 CC quantitating a polypeptide in tissue, as molecular weight markers and as  
 CC a food supplement. (II) and its binding partners are useful in medical  
 CC imaging of sites expressing (II). (I) and (II) are useful for treating  
 CC disorders involving aberrant protein expression or biological activity.  
 CC The polypeptide and polynucleotide sequences have applications in  
 CC diagnostics, forensics, gene mapping, identification of mutations  
 CC responsible for genetic disorders or other traits to assess biodiversity  
 CC and to produce other types of data and products dependent on DNA and  
 CC amino acid sequences. ABG00010-ABG30377 represent novel human  
 CC diagnostic amino acid sequences of the invention.  
 CC Note: The sequence data for this patent did not appear in the printed  
 CC specification, but was obtained in electronic format directly from WIPO  
 CC at ftp.wipo.int/pub/published\_pct\_sequences.  
 XX  
 SQ Sequence 369 AA;

Query Match 77.3%; Score 1463; DB 22; Length 369;  
 Best Local Similarity 75.8%; Pred. No. 9.6e-163;  
 Matches 279; Conservative 43; Mismatches 40; Indels 6; Gaps 5;  
 QY 1 MGLRKNARNPPVLSHEFVQNHADMSCVGMFFVLGLMFECTAEMSIVFLTLQHGVPVP 60  
 DB 1 mglrktstnpvlsqefilqhadiavscvgmffilgfvfegtaeasivfltlqhsavp 60  
 QY 61 -ABGLPSGRTLYHYGVKDLATVFFYMLVAIIHATIQEYVLDKLSRRLQTLKGQKNLN 119  
 DB 61 aaeeqatgsksllyygvkdlatvffymvliiathiqeyvldklnkrmqtkakqknfn 120



PT 830 Primers useful for synthesizing full length cDNA clones and their  
 XX use in genetic manipulation  
 PS Claim 8; SEQ ID NO 2725; 1380pp + sequence listing; English.  
 XX  
 CC The invention relates to primers for synthesizing full length cDNA  
 CC clones. 830 cDNA molecules encoding a human protein have been  
 CC isolated and nucleotide sequences of 5'- and 3'-ends of the cDNA  
 CC molecules have been determined. Primers for synthesizing the full length  
 CC cDNA are useful for clarifying the function of the protein encoded by  
 CC the cDNA. The full length clones were obtained by construction of full  
 CC length enriched cDNA libraries that were synthesised by the oligo-capping  
 CC method. The primers enable the production of the full length cDNA easily  
 CC without any special methods. The present sequence is a polypeptide  
 CC encoded by a full length human cDNA of the invention.  
 CC Note: The sequence data for this patent did not form part of the printed  
 CC specification, but was obtained in CD-ROM format directly from EPO.  
 XX  
 SQ Sequence 369 AA;

Query Match 77.3%; Score 1462; DB 22; Length 369;  
 Best Local Similarity 75.5%; Pred. No. 1.3e-162;  
 Matches 278; Conservative 44; Mismatches 40; Indels 6; Gaps 5;

QY 1 MGLRKKNARNPPVLSHEFMVQNHADMVSCVGMFFVLGLMFEQTAEMSIIVFTLQHGVPVP 60  
 DB 1 mglrkkstknppvlsqefilgnhadvscvgmffilglvfestaesivftldhsavp 60

QY 61 -AGELPSGSTLYHYGVKDLATVFFVLMVAIIHATIQEYVLDKLSRRQLTKGQNKLN 119  
 DB 61 aaeqatgskslvyvgvkdlatvffvmlvaaiihatiqeyvldklnrmqftkqkqkn 120

QY 120 EAGLSVFIYVSGIWMITLASENCLSDPTLLWKSOPHNMTFQMKFFVISOLAYWFHSF 179  
 DB 121 esqgsvfyfscwgtfllisencldspdlwarkphsmmfqkffvsglaywfhaf 180

QY 180 PELYFQVRKQDIPGQIYIGLHFGGAYLLYNLHGLLLMLHYAVELSSVCSLLY 239  
 DB 181 pelyfqtktkqdlprqlvyglhftitgalyilnhlglvllvhyfvelshmcgify 240

QY 240 FGERYQKGLSWPIYFISGRVTLIVSVTVGLHAGT-NRNGNALSGNVNVLAAKTAV 298  
 DB 241 fdekvyqgslwaiflgrlvtlivsvtvglhagsqncpdlagnvnlvaaktav 300

QY 299 LSSSCSIQVYITWTLTVWLQWLEDANLHV-CGRKRRSR-SRKGTEN--GVENPNRIDS 354  
 DB 301 lssscsiqavvtwnliltlwlgrwvdsniqacmkkkrrsskrktengvgvetsnrvc 360

QY 355 PPKKKEKA 362  
 DB 361 ppkrkkes 368

RESULT 7  
 ID AAB43601 standard; Protein; 416 AA.  
 AC AAB43601;  
 XX  
 DT 08-FEB-2001 (first entry)  
 XX  
 DE Human cancer associated protein sequence SEQ ID NO:1046.  
 XX  
 KW Human; cancer associated gene; cancer antigen; detection; cancer;  
 KW diagnosis; cytostatic; proliferative; vulnerable; immunomodulator;  
 KW antiidiabetic; antidiabetic; antirheumatic; antithrombotic; antiviral;  
 KW antiinflammatory; antithyroid; antiallergic; antibacterial; cardiant;  
 KW dermatological; neuroprotective; thrombolytic; coagulant; neutropenic;  
 KW vasotrophic; antipsoriatic; antiangiogenic; gene therapy; inflammation;  
 KW immune disorder; haematopoietic cell disorder; autoimmune disorder;  
 KW allergic reaction; graft versus host disease; organ rejection;  
 KW haemostatic; thrombolytic; cardiovascular disorder; infection;

KW neurological disease; drug screening.  
 XX  
 OS Homo sapiens.  
 XX  
 PN WO200055350-A1.  
 XX  
 PD 21-SEP-2000.  
 XX  
 PF 08-MAR-2000; 2000WO-US05882.  
 XX  
 PR 12-MAR-1999; 9905-0124270.  
 XX  
 PA (HUMA-) HUMAN GENOME SCI INC.  
 XX  
 PI Rosen CA, Ruben SM;  
 XX  
 DR WPI; 2000-587533/55.  
 XX  
 DR N-PSDB; AAC77810.  
 XX  
 PT Novel isolated nucleic acids comprising sequences encoding peptides  
 XX useful for treating or diagnosing e.g. cancer -  
 PT  
 XX Claim 11; Page 1634-1636; 2352pp; English.  
 PS  
 XX AAC77607 to AAC78448 encode the human cancer associated proteins given  
 CC in ABA43398 to ABA44239. The proteins can have activities based on the  
 CC tissues and cells the genes are expressed in. Example of activities  
 CC include: cytostatic; proliferative; vulnerable; immunomodulator;  
 CC antiidiabetic; antidiabetic; antirheumatic; antithyroid;  
 CC antiinflammatory; antithyroid; antipsoriatic; antidiabetic; antiviral;  
 CC dermatological; neuroprotective; cardiant; thrombolytic; coagulant;  
 CC neutropenic; vasotrophic; antipsoriatic and antiangiogenic. The  
 CC polynucleotides and polypeptides can be used for preventing, treating or  
 CC ameliorating medical conditions and diagnosing pathological conditions.  
 CC Polynucleotides, polypeptides, antibodies, agonists and antagonists from  
 CC the present invention may be used to treat immune disorders by activating  
 CC or inhibiting the proliferation, differentiation or mobilisation of  
 CC immune cells, to treat disorders of haematopoietic cells, autoimmune  
 CC disorders, allergic reactions, graft versus host disease and organ  
 CC rejection, modulate haemostatic or thrombolytic activity, modulate  
 CC inflammation, cancers, cardiovascular disorders, neurological disease and  
 CC bacterial or viral infections. The peptides, nucleotides, antibodies,  
 CC agonists and antagonists may be also be used in drug screens. AAC78449 to  
 CC AAC78457 and ABA44240 represent sequences used in the exemplification of  
 CC the present invention.  
 XX  
 SQ Sequence 416 AA;

Query Match 60.0%; Score 1134.5; DB 21; Length 416;  
 Best Local Similarity 58.7%; Pred. No. 4.5e-124;  
 Matches 219; Conservative 61; Mismatches 82; Indels 11; Gaps 4;

QY 1 MGLRKKNARNPPVLSHEFMVQNHADMVSCVGMFFVLGLMFEQTAEMSIIVFTLQHGVPVP 60  
 DB 43 mairkstkstknppvlshefvlgnhadvscvgmffilglvfestaesivftldhsavp 102

QY 61 A-BGLPSGSTLYHYGVKDLATVFFVLMVAIIHATIQEYVLDKLSRRQLTKGQNKLN 119  
 DB 103 ateeqatesvlyyvgvkdlatvffvmlvaaiihatiqeyvldklnrmqftkqkqkn 162

QY 120 EAGLSVFIYVSGIWMITLASENCLSDPTLLWKSOPHNMTFQMKFFVISOLAYWFHSF 179  
 DB 163 esqglsafyfacvgtgftllisencldspdlwarkphsmmfqkffvsglaywfhaf 222

QY 180 PELYFQVRKQDIPGQIYIGLHFGGAYLLYNLHGLLLMLHYAVELSSVCSLLY 239  
 DB 223 pelyfqtktkqdlprqlvyglhftitgalyilnhlglvllvhyfvelshmcgify 282

QY 240 FGERYQKGLSWPIYFISGRVTLIVSVTVGLHAGT-NRNGNALSGNVNVLAAKTAV 298  
 DB 283 fsnekvyqgslwaiflgrlvtlivsvtvglhagsqncpdlagnvnlvaaktav 342

QY 299 LSSCSIQYIYIWTTLTTLVWLQWLEADANLHVCGRRKRR-----SRSRKGTEGVE---NP 349  
 Db 343 lasicvtqafmmwkkfinfglrwrhsafapavkkkptvtckgr:skkgtengvngtlts 402  
 QY 350 NRIDSPPKKKEKA 362  
 Db 403 nvadsprnkkks 415

RESULT 8  
 AAY48434  
 ID AAY48434 standard; Protein; 304 AA.  
 AC AAY48434;  
 XX  
 DT 08-DEC-1999 (first entry)  
 DE Human prostate cancer-associated protein 131.  
 KW Expressed sequence tag; EST; prostate; tumor; treatment; gene therapy;  
 KW cancer; tissue specificity; human.  
 XX  
 OS Homo sapiens.  
 XX  
 PN DE19811194-A1.  
 PD 16-SEP-1999.  
 PF 10-MAR-1998; 98DE-1011194.  
 XX 10-MAR-1998; 98DE-1011194.  
 PR  
 XX (META-) METAGEN GES GENOMFORSCHUNG MBH.  
 PA Specht T, Hinzmann B, Schmitt A, Pilarsky C, Dahl E, Rosenthal A;  
 PI WPI; 1999-519629/44.  
 DR N-PSDB; AAZ33531.  
 XX  
 PT New nucleic acid expressed at high level in normal prostatic tissue and  
 PT encoded polypeptides, used to treat cancer and screen for therapeutic  
 PT agents  
 XX  
 PS Claim 25; 174; 194pp; German.  
 XX  
 CC This invention describes novel nucleic acid sequences (A) that are  
 CC expressed at high level in normal prostatic tissue. Polypeptides (I)  
 CC encoded by (A) are used: (a) for identifying agents for treatment of  
 CC prostatic cancer and (b) for therapy of prostate cancer, optionally  
 CC where expressed by gene therapy methods. (A) is also used to isolate  
 CC full-length genes (for gene therapy) and for recombinant production of  
 CC (I), which can be used to raise specific antibodies. (A) are identified  
 CC by assembly of ESTs (expressed sequence tags) before these are analyzed  
 CC for expression pattern (tissue specificity). This approach eliminates  
 CC many of the false results, as regards tissue specificity, associated  
 CC with known methods that use single (usually short) ESTs. AAY48304-Y48456  
 CC represent peptides encoded by the expressed sequence tags described in  
 CC the method of the invention.  
 XX  
 SQ Sequence 304 AA;

Query Match 48.78; Score 922; DB 20; Length 304;  
 Best Local Similarity 58.18; Pred. No. 2.5e-99;  
 Matches 176; Conservative 48; Mismatches 69; Indels 10; Gaps 3;

QY 70 TLHYGVKDLATVFYMLVAIIIIHATIQEYVLDKLSRRLOLTGKQNKLEAGSLQVFIY 129  
 Db 1 sllyygyrdlatvfymlyvailhavqeymldkinrmhfstkskfnesqqlsafyl 60

QY 130 VSGIWMIIASENCLSDPTLLWKSQPHNMVTFOMKFFIYSLQAYWPHSPPELYFQVRK 189  
 Db 61 facvwtgftllisenysdptllwrayphnltmtgmkfysqlaywlhafpelyfqtktk 120

QY 190 QDIPQLIYIGLHFLPHIGGAYLLYLNLHGLLLMLHYAVELLSSVCSLLYFGDERYQKGL 249  
 Db 121 edipqlyyiglyfhiagaylnlnhlglylvlyhvfeflhisrlfyfsnekykqgf 180  
 QY 250 SLWPIVFIISGRLLVTLIVSVTVGLHLA-GTNRNGNALSGNVNLAAKIAYLSSSCSIQVY 308  
 Db 181 slwavflvrltllslvtlvgfglaraenqkldfatgnfnvlavriavlasivtqaf 240  
 QY 309 ITWTLTTLVWLQWLEADANLHVCGRRKRR-----SRSRKGTEGVE---NPNRIDSPPKKK 359  
 Db 241 mmwkkfinfglrwrhsafapavkkkptvtckgrskkgtengvngtltsnvadsprnkk 300  
 QY 360 EKA 362  
 Db 301 eks 303

RESULT 9  
 AAM95708  
 ID AAM95708 standard; Protein; 384 AA.  
 XX  
 AC AAM95708;  
 XX  
 DT 21-NOV-2001 (first entry)  
 DE Human reproductive system related antigen SEQ ID NO: 4366.  
 XX  
 KW Human; reproductive system related antigen; reproductive system disorder;  
 KW cancer; gene therapy.  
 XX  
 OS Homo sapiens.  
 XX  
 PN WO200155320-A2.  
 PD 02-AUG-2001.  
 XX  
 PF 17-JAN-2001; 2001WO-US01339.  
 XX  
 PR 31-JAN-2000; 2000US-0179065.  
 PR 04-FEB-2000; 2000US-0180628.  
 PR 24-FEB-2000; 2000US-0184664.  
 PR 02-MAR-2000; 2000US-0186350.  
 PR 16-MAR-2000; 2000US-0189874.  
 PR 17-MAR-2000; 2000US-0190076.  
 PR 18-APR-2000; 2000US-0198123.  
 PR 19-MAY-2000; 2000US-0205515.  
 PR 07-JUN-2000; 2000US-0209467.  
 PR 28-JUN-2000; 2000US-0214886.  
 PR 30-JUN-2000; 2000US-0215135.  
 PR 07-JUL-2000; 2000US-0216647.  
 PR 07-JUL-2000; 2000US-0216880.  
 PR 11-JUL-2000; 2000US-0217487.  
 PR 11-JUL-2000; 2000US-0217496.  
 PR 14-JUL-2000; 2000US-0218290.  
 PR 26-JUL-2000; 2000US-0220963.  
 PR 26-JUL-2000; 2000US-0220964.  
 PR 14-AUG-2000; 2000US-0224518.  
 PR 14-AUG-2000; 2000US-0224519.  
 PR 14-AUG-2000; 2000US-0225213.  
 PR 14-AUG-2000; 2000US-0225214.  
 PR 14-AUG-2000; 2000US-0225266.  
 PR 14-AUG-2000; 2000US-0225267.  
 PR 14-AUG-2000; 2000US-0225268.  
 PR 14-AUG-2000; 2000US-0225270.  
 PR 14-AUG-2000; 2000US-0225447.  
 PR 14-AUG-2000; 2000US-0225757.  
 PR 14-AUG-2000; 2000US-0225758.  
 PR 14-AUG-2000; 2000US-0225759.  
 PR 18-AUG-2000; 2000US-0226279.  
 PR 22-AUG-2000; 2000US-0226681.  
 PR 22-AUG-2000; 2000US-0226868.  
 PR 22-AUG-2000; 2000US-0227182.







Query Match 29.9%; Score 566; DB 22; Length 368;  
Best Local Similarity 37.0%; Pred. No. 2.2e-57;  
Matches 136; Conservative 64; Mismatches 130; Indels 38; Gaps 9;

QY 4 RKKRNAPPVLSHEFMVONHADMVSCVGMFFVGLMFEPTAEMSIYFLTLQHGCV--VWPA 61  
DB 9 rktknknpplshsfvignhadliiscvamfvvglnmestaafasafisllhnnvsgebps 68  
QY 62 EGLPSGSRITLYHYGVKDLATVFFVYMLVAIIHATIQEYVLDKLSRRLQLTGKQKNLNEA 121  
DB 69 reqpygkpytiagikdycaiffytltciimhaiiqefvldkiskkhskfklarfnes 128  
QY 122 GQLSVFVIVSGINGMIILASENCLSDPTLLWKSQPHNMFTQMKFFYISQLAYWFHSFPE 181  
DB 129 gqlvafyllsfvghavllkegylgvqagwefpdpmpsfllhkfvyvqglayylhmlpe 188  
QY 182 LYFQKVR-KODIPGOLYIGLHFLPHIGG-----AYLLYLNHLGLLLMLHYAVELLSSV 234  
DB 189 lyfqtktkeeqpkivh-----sigsftllvlaytlfqrlalvllthysellshv 242  
QY 235 CSLL--YFGDERYOKGLSLMPVIFISGRVTLIVSVVTVGLHLAGTNRNGNALSGNVNVL 292  
DB 243 fqligvfdreerlaklrnnnavflirfatsvigtlyyggv-rslalggli--- 298  
QY 293 AAKTAVLSSCSQIVYITWTTLTWLQRLWLEDANLHVCGRRSRKRGKTENGVPNRI 352  
DB 299 -----alqgylvfsfiteqlrakreakk-----eakreakialqtktkpktpk-- 341  
QY 353 DSPPKKKE 360  
DB 342 dkvrkke 349

RESULT 12  
ABB67362  
ID ABB67362 standard; Protein; 1575 AA.  
XX AC ABB67362;  
XX DT 26-MAR-2002 (first entry)  
XX DE Drosophila melanogaster polypeptide SEQ ID NO 28878.  
XX KW Drosophila; developmental biology; cell signalling; insecticide;  
XX KW pharmaceutical.  
XX OS Drosophila melanogaster.  
XX PN WO200171042-A2.  
XX PD 27-SEP-2001.  
XX PF 23-MAR-2001; 2001WO-US09231.  
XX PR 23-MAR-2000; 2000US-191637P.  
XX PR 11-JUL-2000; 2000US-0614150.  
XX PA (PEKE ) PE CORP NY.  
XX PI Venter JC, Adams M, Li PWD, Myers EW;  
XX DR WPI; 2001-656860/75.  
XX DR N-PSDB; ABL11465.  
XX PT New isolated nucleic acid detection reagent for detecting 1000 or more  
XX PT genes from Drosophila and for elucidating cell signalling and cell-cell  
XX PT interactions -  
XX PS Disclosure; SEQ ID NO 28878; 21pp + Sequence Listing; English.  
XX CC The invention relates to an isolated nucleic acid detection reagent  
XX CC capable of detecting 1000 or more genes from Drosophila. The invention is

useful in developmental biology and in elucidating cell signalling and  
cell-cell interactions in higher eukaryotes for the development of  
insecticides, therapeutics and pharmaceutical drugs. The invention  
discloses genomic DNA sequences (ABL16176-ABL30511), expressed DNA  
sequences (ABL01840-ABL16175) and the encoded proteins  
(ABB57737-ABB72072).  
The sequence data for this patent did not form part of the printed  
CC specification, but was obtained in electronic format directly from WIPO  
CC at ftp.wipo.int/pub/published\_pct\_sequences.  
XX Sequence 1575 AA;  
SQ

Query Match 29.9%; Score 566; DB 22; Length 1575;  
Best Local Similarity 37.0%; Pred. No. 2e-56;  
Matches 136; Conservative 64; Mismatches 130; Indels 38; Gaps 9;

QY 4 RKKRNAPPVLSHEFMVONHADMVSCVGMFFVGLMFEPTAEMSIYFLTLQHGCV--VWPA 61  
DB 1216 rktknknpplshsfvignhadliiscvamfvvglnmestaafasafisllhnnvsgebps 1275  
QY 62 EGLPSGSRITLYHYGVKDLATVFFVYMLVAIIHATIQEYVLDKLSRRLQLTGKQKNLNEA 121  
DB 1276 reqpygkpytiagikdycaiffytltciimhaiiqefvldkiskkhskfklarfnes 1335  
QY 122 GQLSVFVIVSGINGMIILASENCLSDPTLLWKSQPHNMFTQMKFFYISQLAYWFHSFPE 181  
DB 1336 gqlvafyllsfvghavllkegylgvqagwefpdpmpsfllhkfvyvqglayylhmlpe 1395  
QY 182 LYFQKVR-KODIPGOLYIGLHFLPHIGG-----AYLLYLNHLGLLLMLHYAVELLSSV 234  
DB 1396 lyfqtktkeeqpkivh-----sigsftllvlaytlfqrlalvllthysellshv 1449  
QY 235 CSLL--YFGDERYOKGLSLMPVIFISGRVTLIVSVVTVGLHLAGTNRNGNALSGNVNVL 292  
DB 1450 fqligvfdreerlaklrnnnavflirfatsvigtlyyggv-rslalggli--- 1505  
QY 293 AAKTAVLSSCSQIVYITWTTLTWLQRLWLEDANLHVCGRRSRKRGKTENGVPNRI 352  
DB 1506 -----alqgylvfsfiteqlrakreakk-----eakreakialqtktkpktpk-- 1548  
QY 353 DSPPKKKE 360  
DB 1549 dkvrkke 1556

RESULT 13  
AAG00189  
ID AAG00189 standard; Protein; 125 AA.  
XX AC AAG00189;  
XX DT 06-OCT-2000 (first entry)  
XX DE Human secreted protein, SEQ ID NO: 4270.  
XX KW Human; 5' EST; expressed sequence tag; secreted protein; cDNA isolation;  
XX KW gene therapy; chromosome mapping.  
XX OS Homo sapiens.  
XX PN EP1033401-A2.  
XX PD 06-SEP-2000.  
XX PF 21-FEB-2000; 2000EP-0200610.  
XX PR 26-FEB-1999; 99US-0122487.  
XX PR (GEST ) GENSET.  
XX PI Dumas Milne Edwards J, Duclert A, Giordano J;  
XX

DR WPI: 2000-500381/45.  
 DR N-PSDB; AAC00195.  
 XX  
 PT New nucleic acid that is a 5' expressed sequence tag (5' EST) for  
 PT obtaining cDNAs and genomic DNAs that correspond to 5'ESTs and for  
 PT diagnostic, forensic, gene therapy and chromosome mapping procedures -  
 XX  
 PS Claim 13; SEQ ID 4270; 71pp + CD-ROM; English.  
 XX  
 CC The present sequence is a polypeptide encoded by one of a large number  
 CC of 5' ESTs derived from mRNAs encoding secreted proteins. The 5' ESTs  
 CC were prepared from total human RNAs or polyA+ RNAs derived from 30  
 CC different tissues. EST sequences usually correspond mainly to the 3'  
 CC untranslated region (UTR) of the mRNA because they are often obtained  
 CC from oligo-dT primed cDNA libraries. Such ESTs are not well suited for  
 CC isolating cDNA sequences derived from the 5' ends of mRNAs and even in  
 CC those cases where longer cDNA sequences have been obtained, the full 5'  
 CC UTR is rarely included. 5' ESTs are derived from mRNAs with intact 5'  
 CC ends and can therefore be used to obtain full length cDNAs and genomic  
 CC DNAs. 5' ESTs are also used in diagnostic, forensic, gene therapy and  
 CC chromosome mapping procedures. They are used to obtain upstream  
 CC regulatory sequences and to design expression and secretion vectors.  
 XX  
 SQ Sequence 125 AA;

Query Match 21.9%; Score 414.5; DB 21; Length 125;  
 Best Local Similarity 64.0%; Pred. No. 2.8e-40;  
 Matches 80; Conservative 24; Mismatches 20; Indels 1; Gaps 1;  
 QY 1 MGLRKNRNPVLSHEFMVONHADMVSCVGMFFVLGLMFEGETAEMSVIFLTLOHGVVVP 60  
 Db 1 mairkstkppvlshfvlghadivscvamlglmlfmitakasiifvldgynvtp 60  
 QY 61 A-EGLPSSGRTLYHYGVKDLATVFFYMLVAIIHATIQEYVLDKLSRLOLTCKQKQKLN 119  
 Db 61 ateeqatesvslyyygiklatvffymvlaiihaviqeymldknrmxfsktxhskfn 120  
 QY 120 EAGQL 124  
 Db 121 esgql 125

RESULT 14  
 AAR86810  
 ID AAR86810 standard; Protein; 411 AA.

XX AAR86810;

XX 28-MAY-1997 (first entry)

DE Saccharomyces cerevisiae LAG1 protein.

XX Saccharomyces cerevisiae; LAG1; life-span limiting domain;  
 KW life-span extending domain; stress tolerance; longevity;  
 KW recombinant protein production.

OS Saccharomyces cerevisiae.

XX Key Location/Qualifiers

FT Domain 1..190

FT Domain /note= "life-span limiting domain"

FT Domain 194..411

FT Domain /note= "life-span extending domain"

XX W09533834-AL.

XX 14-DEC-1995.

XX 02-JUN-1995; 95WO-US06725.

XX 08-NOV-1994; 94US-0336031.

PR 03-JUN-1994; 94US-0253875.

XX (RESE ) RESEARCH CORP TECHNOLOGIES INC.  
 XX Jazwinski SM;  
 XX WPI: 1996-040238/04.  
 DR N-PSDB; AAT07263.  
 XX  
 PT Eukaryotic LAG1 gene and protein - controls longevity, stress  
 PT tolerance and reproductive capacity of eukaryotic cells, for  
 PT improved prodn. of recombinant proteins.  
 XX  
 PS Claim 23; Page 100-102; 154pp; English.  
 XX  
 CC The Saccharomyces cerevisiae LAG1 gene product is composed of a  
 CC life-span limiting domain (see AAR86812) and a life-span extending  
 CC domain (see AAR86811). Overexpression of the LAG1 gene in older  
 CC cells has a rejuvenating effect, which not only increases  
 CC cellular life span, but also reproductive capacity and cellular  
 CC tolerance to stress factors such as starvation and low pH. These  
 CC cells may be used for the production of recombinant proteins. By  
 CC increasing the life span of recombinant cells, the need for  
 CC overexpression of recombinant gene products is avoided, and  
 CC therefore any subsequent adverse effects on the host cell.  
 XX  
 SQ Sequence 411 AA;

Query Match 7.8%; Score 148.5; DB 17; Length 411;  
 Best Local Similarity 22.0%; Pred. No. 2.9e-08;  
 Matches 68; Conservative 46; Mismatches 106; Indels 89; Gaps 11;  
 QY 26 MYSCVGMFFVLGLMFEGETAEMSVIFLTLOHGVVVPAGELPSGSRTRYHYGVKDLATVFFY 85  
 Db 91 lvcvysavflsgnrtesplhmfaivsq-----vsgtds-----yakgikdisfvyf 139  
 QY 86 MLVAIIIIHATIQEYVLDKLSRRLQLTCKQKQKLNAGOLSVFYI-VSGINGMILL-ASEN 143  
 Db 140 mifftflreflmdvvrftvylntvsehrgkrmlegmaifcvgsgpfglymyhsdl 199  
 QY 144 CLSDPTLLKWSQPHNMMTFMKFFVISOLAYW-----FHSFPE 181  
 Db 200 wlfktkpmtytpvntnplfkfifygqaafwagqacvllqlckprkdykelyfhvht 259  
 QY 182 L-----YFQKVRKODIPQOLIYI-----GLHLFH 205  
 Db 260 llllwsyyvfhtkm-----glaiytdmvsdfllsktlnylnsvftpfvfglfvf- 312  
 QY 206 IGGAYLLYLNHLGLLLMLHYAVELLSSVCSLLYEGDERYKGLSLMPYFISGRVLTLI 265  
 Db 313 -----fwylrhvnrilrllswvltefrhagnyvlmfatqgkckwisl-pivfv-----ll 361  
 QY 266 VSVVTVGLH 274  
 Db 362 aalqlvnly 370

RESULT 15

AAAY00876

ID AAAY00876 standard; Protein; 394 AA.

XX AAAY00876;

DT 21-MAY-1999 (first entry)

XX Human LAPH-1 protein sequence.

XX Longevity-assurance protein homologue; LAPH-1; human; cell proliferation;  
 KW longevity-assurance protein homologue; LAPH-2; signal transduction;  
 KW cell cycle regulation; apoptosis; cellular homeostatic pathway; aging;  
 KW cancer; inflammation; autoimmune disease; infection;  
 XX neurodegenerative disorder.





GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 20, 2002, 14:34:08 ; Search time 13.07 Seconds  
(without alignments)  
678.385 Million cell updates/sec

Title: US-09-807-470-2

Perfect score: 1892

Sequence: 1 MGLRKKARPPVLSHEFMV.....NGVENPNRIDSPPKKKEKAP 363

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 231628 seqs, 24425594 residues

Total number of hits satisfying chosen parameters: 231628

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

- 1: /cgn2.6/ptodata/2/1aa/5A\_COMB.pap:\*
- 2: /cgn2.6/ptodata/2/1aa/5B\_COMB.pap:\*
- 3: /cgn2.6/ptodata/2/1aa/6A\_COMB.pap:\*
- 4: /cgn2.6/ptodata/2/1aa/6B\_COMB.pap:\*
- 5: /cgn2.6/ptodata/2/1aa/PCTUS\_COMB.pap:\*
- 6: /cgn2.6/ptodata/2/1aa/backfiles1.pap:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	148.5	7.8	411	2	US-08-336-031-2
2	148.5	7.8	411	2	US-08-902-853-7
3	148.5	7.8	411	5	PCT-US95-06725-2
4	138.5	7.3	394	2	US-08-902-853-1
5	125	6.6	387	2	US-08-902-853-6
6	110	5.8	380	2	US-08-902-853-3
7	89	4.7	649	4	US-08-800-291B-5
8	89	4.7	649	4	US-08-800-291B-6
9	89	4.7	650	4	US-08-800-291B-4
10	87	4.6	1476	4	US-09-256-703-2
11	87	4.6	1479	2	US-08-951-912-4
12	87	4.6	1479	4	US-09-174-077-4
13	87	4.6	1480	1	US-07-637-621-2
14	87	4.6	1480	1	US-08-136-742A-2
15	87	4.6	1480	1	US-08-135-809A-2
16	87	4.6	1480	1	US-08-466-886-17
17	87	4.6	1480	2	US-08-951-912-2
18	87	4.6	1480	2	US-08-951-912-6
19	87	4.6	1480	2	US-08-469-461-4
20	87	4.6	1480	2	US-08-469-461-4
21	87	4.6	1480	2	US-08-691-605-2
22	87	4.6	1480	2	US-08-453-552A-14
23	87	4.6	1480	3	US-07-890-609-2
24	87	4.6	1480	3	US-07-890-609-4
25	87	4.6	1480	3	US-09-248-026-2
26	87	4.6	1480	4	US-08-469-617-17
27	87	4.6	1480	4	US-08-681-838A-2

28	87	4.6	1480	4	US-08-681-838A-3	Sequence 3, Appli
29	87	4.6	1480	4	US-09-174-077-2	Sequence 2, Appli
30	87	4.6	1480	4	US-09-174-077-6	Sequence 6, Appli
31	87	4.6	1480	5	PCT-US93-11667-2	Sequence 2, Appli
32	87	4.6	1480	6	5240846-5	Patent No. 5240846
33	85	4.5	193	2	US-08-336-031-6	Sequence 6, Appli
34	85	4.5	193	5	PCT-US95-06725-6	Sequence 6, Appli
35	82	4.3	2183	3	US-08-746-111-5	Sequence 5, Appli
36	80	4.2	617	1	US-07-879-617A-11	Sequence 11, Appli
37	80	4.2	617	1	US-08-301-722A-3	Sequence 3, Appli
38	80	4.2	617	1	US-08-240-783B-3	Sequence 3, Appli
39	80	4.2	617	1	US-08-753-985-11	Sequence 11, Appli
40	80	4.2	617	3	US-09-084-813-3	Sequence 3, Appli
41	80	4.2	617	5	PCT-US92-09662-3	Sequence 2, Appli
42	79	4.2	326	4	US-08-986-768-2	Sequence 2, Appli
43	79	4.2	602	1	US-08-295-814E-2	Sequence 2, Appli
44	79	4.2	602	4	US-09-343-361-2	Sequence 2, Appli
45	79	4.2	602	5	PCT-US93-01959-2	Sequence 2, Appli

## ALIGNMENTS

RESULT 1  
US-08-336-031-2  
; Sequence 2, Application US/08336031  
; Patent No. 5817782  
; GENERAL INFORMATION:  
; APPLICANT: Jazwinski, S. M.  
; TITLE OF INVENTION: LAG1: A GENE FOR INCREASING THE  
; LONGEVITY OF EUKARYOTES  
; NUMBER OF SEQUENCES: 7  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Scully, Scott, Murphy & Presser  
; STREET: 400 Garden City Plaza  
; CITY: Garden City  
; STATE: New York  
; COUNTRY: United States  
; ZIP: 11530  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent In Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/336,031  
; FILING DATE:  
; CLASSIFICATION: 514  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/253,875  
; FILING DATE: 03-JUN-1994  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Digiglio, Frank S.  
; REGISTRATION NUMBER: 31,346  
; REFERENCE/DOCKET NUMBER: 9303Z  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (516) 742-4343  
; TELEFAX: (516) 742-4366  
; TELEX: 230 901 SANS UR  
; INFORMATION FOR SEQ ID NO: 2:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 411 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
US-08-336-031-2

Query Match 7.8%; Score 148.5; DB 2; Length 411;  
Best Local Similarity 22.0%; Pred. No. 1.1e-08;  
Matches 68; Conservative 46; Mismatches 106; Indels 89; Gaps 11;

QY 26 MVSVCGMFFVLGMFEGTMSIVFLTQHGVVVPAEGLPGSRTLYHGVKDLATVFFY 85

```
Db 91 LVCVTSATFLSNGRTESNPLMFVAISQ-----VDGTD-----YAKGIKDLSFVFF 139
QY 86 MLVAIIIIHATIQEYVLDKLSRRQLTKGKQNKLEAGOLSVFYI-VSGIWMGMIIL-ASEN 143
Db 140 MIFFTFLREFLMDVIRPFTVYLVNTSEHROKRMLEQMYAIFYCGVSPGFLYIMYHSDL 199
QY 144 CLSDPTLLWKSPHNMFTQMKFFVVISQLAYW-----FHSFPE 181
Db 200 WLFTKPMRTYPTVITNPFLEKFIYLGQAFMAOQACVLVLQLEKPKDYKELVFHHIVT 259
QY 182 L-----YFQVKRKODIPGOLIYI-----GLAIYITMDVSDFFLSKLTNLYNSVFTPFVGLFVF- 312
Db 260 LLLIWSVVFHFTKM-----GLAIYITMDVSDFFLSKLTNLYNSVFTPFVGLFVF- 312
QY 206 ICGAYLLYLNLHGLLLMLHYAVELLSSVCSLLYFGDERYOKGSLMPTVIFISGRVTLI 265
Db 313 -----FWIYLRHVNIIRILWSVLTEFRHGNVYLNFAATQYKCMWISL-PIVFFV-----LI 361
QY 266 VSVTVTVGLH 274
Db 362 AALQLVNLV 370

RESULT 2
US-08-902-853-7
; Sequence 7, Application US/08902853
; Patent No. 5945330
; GENERAL INFORMATION:
; APPLICANT: Hillman, Jennifer L.
; APPLICANT: Corley, Neil C.
; APPLICANT: Shah, Purvi
; APPLICANT: Lal, Preeti
; TITLE OF INVENTION: HUMAN LONGEVITY-ASSURANCE PROTEIN HOMOLOGS
; NUMBER OF SEQUENCES: 7
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Incyte Pharmaceuticals, Inc.
; STREET: 3174 Porter Drive
; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA
; ZIP: 94304
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSEQ for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/902,853
; FILING DATE: Herewith
; CLASSIFICATION: ?
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Billings, Lucy J.
; REGISTRATION NUMBER: 36,749
; REFERENCE/DOCKET NUMBER: PF-0345 US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415-855-0555
; TELEFAX: 415-845-4166
; TELEX:
; INFORMATION FOR SEQ ID NO: 7:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 411 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; IMMEDIATE SOURCE:
; LIBRARY: Genbank
; CLONE: 541568
; US-08-902-853-7
```

```
Query Match 7.8%; Score 148.5; DB 2; Length 411;
Best Local Similarity 22.0%; Pred. No. 1.le-08;
Matches 68; Conservative 46; Mismatches 106; Indels 89; Gaps 11;

QY 26 MYSCVGMFEVLGLMEFGTAEMSIYVFLTLQHGVVVPAEGLSPGSRTLHYGVKDLATVFFY 85
Db 91 LVCVTSATFLSNGRTESNPLMFVAISQ-----VDGTD-----YAKGIKDLSFVFF 139
QY 86 MLVAIIIIHATIQEYVLDKLSRRQLTKGKQNKLEAGOLSVFYI-VSGIWMGMIIL-ASEN 143
Db 140 MIFFTFLREFLMDVIRPFTVYLVNTSEHROKRMLEQMYAIFYCGVSPGFLYIMYHSDL 199
QY 144 CLSDPTLLWKSPHNMFTQMKFFVVISQLAYW-----FHSFPE 181
Db 200 WLFTKPMRTYPTVITNPFLEKFIYLGQAFMAOQACVLVLQLEKPKDYKELVFHHIVT 259
QY 182 L-----YFQVKRKODIPGOLIYI-----GLAIYITMDVSDFFLSKLTNLYNSVFTPFVGLFVF- 312
Db 260 LLLIWSVVFHFTKM-----GLAIYITMDVSDFFLSKLTNLYNSVFTPFVGLFVF- 312
QY 206 ICGAYLLYLNLHGLLLMLHYAVELLSSVCSLLYFGDERYOKGSLMPTVIFISGRVTLI 265
Db 313 -----FWIYLRHVNIIRILWSVLTEFRHGNVYLNFAATQYKCMWISL-PIVFFV-----LI 361
QY 266 VSVTVTVGLH 274
Db 362 AALQLVNLV 370

RESULT 3
PCT-US95-06725-2
; Sequence 2, Application PC/TUS9506725
; GENERAL INFORMATION:
; APPLICANT:
; TITLE OF INVENTION: LAG1: A GENE FOR INCREASING THE
; TITLE OF INVENTION: LONGEVITY OF EUKARYOTES
; NUMBER OF SEQUENCES: 7
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Scully, Scott, Murphy & Presser
; STREET: 400 Garden City Plaza
; CITY: Garden City
; STATE: New York
; COUNTRY: United States
; ZIP: 11530
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US95/06725
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/253,875 & 08/336,031
; FILING DATE: 03-JUN-1994 & 08-NOV-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Digiglio, Frank S.
; REGISTRATION NUMBER: 31,346
; REFERENCE/DOCKET NUMBER: 93032
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (516) 742-4343
; TELEFAX: (516) 742-4366
; TELEX: 230 901 SANS UR
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 411 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; PCT-US95-06725-2
```

Query Match 7.8%; Score 148.5; DB 5; Length 411;  
Best Local Similarity 22.0%; Pred. No. 1.1e-08;  
Matches 68; Conservative 46; Mismatches 106; Indels 89; Gaps 11;  
Qy 26 MVSVCVGFVLLMREGTAEMSVFLDLQHGUVVPAEGLPSGSGRTLYHYGVKDLAVFFY 85  
Db 91 LVCVSYAFVLSGNTESNPLHMFVAISYQ-----VDGTD-----YAKGIKDLSEVFFY 139  
Qy 86 MLVAIIITHATQEVYVLDKLSRRQLTKGKQKLNKAGQSVFYI-VSGIWMIL-ASEN 143  
Db 140 MIEFTLREFLMDVIRPFTVYLVNTEHQRKMLQOMYAFYCVSGPGLYIMVHSD 199  
Qy 144 CLSDPTLLKQSNHMTFQKFFYISQLAYW-----FHSFPE 181  
Db 200 WLFKTKMYRTYPVITNPFELKIFYLQAAFAWAQACVVLQLEKPKDKYELVPHIHT 259  
Qy 182 L-----YFQVKRKODIPGOLIIYI-----GLHLFH 205  
Db 260 LLLIWSVVVFHTM-----GLAIYITMDVSDFFLSLTKLNLNSVTFPFVGLFVF- 312  
Qy 206 IGGAYLYLNHLGMLLMLAYAVELLSSVCSLLYFGDERYQKGLSWPIVIFISGRVLT 265  
Db 313 ----FWIYLRHVNRILWSVLTEFRHGNVNLNFAQOQKWL-SL-PIVFN-----LI 361  
Qy 266 VSVVTVGLH 274  
Db 362 AALQVLNLY 370

RESULT 4  
US-08-902-853-1  
; Sequence 1, Application US/08902853  
; Patent No. 5945330  
; GENERAL INFORMATION:  
; APPLICANT: Hillman, Jennifer L.  
; APPLICANT: Corley, Neil C.  
; APPLICANT: Shah, Purvi  
; APPLICANT: Lal, Preeti  
; TITLE OF INVENTION: HUMAN LONGEVITY-ASSURANCE PROTEIN HOMOLOGS  
; NUMBER OF SEQUENCES: 7  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Incyte Pharmaceuticals, Inc.  
; STREET: 3174 Porter Drive  
; CITY: Palo Alto  
; STATE: CA  
; COUNTRY: USA  
; ZIP: 94304  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Diskette  
; COMPUTER: IBM Compatible  
; OPERATING SYSTEM: DOS  
; SOFTWARE: FastSeq for Windows Version 2.0  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/902,853  
; FILING DATE: Herewith  
; CLASSIFICATION: ?  
; PRIOR APPLICATION NUMBER:  
; FILING DATE:  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Billings, Lucy J.  
; REGISTRATION NUMBER: 36,749  
; REFERENCE/DOCKET NUMBER: PF-0345 US  
; TELEPHONE: 415-855-0555  
; TELEFAX: 415-845-4166  
; TELEX:  
; INFORMATION FOR SEQ ID NO: 1:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 394 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single

TOPOLOGY: linear  
IMMEDIATE SOURCE:  
LIBRARY: LIVERTUT04  
CLONE: 2516821  
US-08-902-853-1  
Query Match 7.3%; Score 138.5; DB 2; Length 394;  
Best Local Similarity 24.7%; Pred. No. 1.5e-07;  
Matches 44; Conservative 42; Mismatches 81; Indels 11; Gaps 4;  
Qy 107 RLQITKQKQKLNKAGQSVFYIVSGIWMILASENCLSDPTLLKQSNHMTFQMKF 166  
Db 126 RPQUTK-----KFCASNRFLFYLSFFVGLSVLHESWLVAPVCMWDYRPNQTLKPSLYW 181  
Qy 167 FYISQLAYWPHSFPELYFQVKRKODIPGQIYIIGLHFIHGAYLLYLNHLGMLLMLHY 226  
Db 182 WYLLLELGYLLSLRLFPD-VKRDKEQVHHFVAVILMTFSYANLLRIGSLVLLHD 240  
Qy 227 AVELLSVCSLLYFGDERYQKGLS-----LWPIVFISGRVLTIVSVTVGLHAGTNR 280  
Db 241 SSDYLLEACKMNY--MQYQQVCDALFISFVFYTRLVLPFTQILYTTYYSISNR 296  
RESULT 5  
US-08-902-853-6  
; Sequence 6, Application US/08902853  
; Patent No. 5945330  
; GENERAL INFORMATION:  
; APPLICANT: Hillman, Jennifer L.  
; APPLICANT: Corley, Neil C.  
; APPLICANT: Shah, Purvi  
; APPLICANT: Lal, Preeti  
; TITLE OF INVENTION: HUMAN LONGEVITY-ASSURANCE PROTEIN HOMOLOGS  
; NUMBER OF SEQUENCES: 7  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Incyte Pharmaceuticals, Inc.  
; STREET: 3174 Porter Drive  
; CITY: Palo Alto  
; STATE: CA  
; COUNTRY: USA  
; ZIP: 94304  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Diskette  
; COMPUTER: IBM Compatible  
; OPERATING SYSTEM: DOS  
; SOFTWARE: FastSeq for Windows Version 2.0  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/902,853  
; FILING DATE: Herewith  
; CLASSIFICATION: ?  
; PRIOR APPLICATION NUMBER:  
; FILING DATE:  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Billings, Lucy J.  
; REGISTRATION NUMBER: 36,749  
; REFERENCE/DOCKET NUMBER: PF-0345 US  
; TELEPHONE: 415-855-0555  
; TELEFAX: 415-845-4166  
; TELEX:  
; INFORMATION FOR SEQ ID NO: 6:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 387 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; ORIGINAL SOURCE:  
; HAPLOTYPED: GenBank  
; IMMEDIATE SOURCE:  
; LIBRARY: 1675382  
; US-08-902-853-6



US-08-902-853-3

Db		325	LIRPYLADMTL---	SEVHVVTGGYATIAGSLGAYISFGIDRTSLINASVMAAPCALAL	381	
Qy		175	WPHSPPELYFQVRKRDIPGQLYIGUHLHFHIGGAYLLYNHLGLLL---	LMHYAVELL	231	
Db		382	SKLVYVEESKFREE-----	GVLTYGDAQNLTEAASTGAATSVKVYVANIAANLI	433	
Qy		232	SSV-----	CSLYFGDERYQRGLS-----	LMPYVFISG	259
Db		434	AFLAVLDFTNAALSGMLGDWDIOGLSFOLICSYILRPFVAFLMG	476		

RESULT 9

```
,
, GENERAL INFORMATION:
, APPLICANT: J.D. Young & C.E. Cass
, TITLE OF INVENTION: CDNA ENCODING NUCLEOSIDE TRANSPORTER
, NUMBER OF SEQUENCES: 8
, CORRESPONDENCE ADDRESS:
, ADDRESSEE: Fish & Richardson P.C.
, STREET: 4225 Executive Square, Suite 1400
, CITY: La Jolla
, STATE: CA
, COUNTRY: USA
, ZIP: 92037
, COMPUTER READABLE FORM:
, MEDIUM TYPE: Floppy disk
, COMPUTER: IBM PC compatible
, OPERATING SYSTEM: PC-DOS/MS-DOS
, SOFTWARE: PatentIn Release #1.0, Version #1.30
, CURRENT APPLICATION DATA:
, APPLICATION NUMBER: US/08/800,291B
, FILING DATE: 13-FEB-1997
, CLASSIFICATION: 424
,
```

```

CURRENT APPLICATION DATA:
APPLICATION NUMBER: US 08/800,291B
FILING DATE: 13-FEB-1997
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/499,314
FILING DATE: 7-JULY-1995
ATTORNEY/AGENT INFORMATION:
NAME: Haile, Lisa A.
REGISTRATION NUMBER: 38,347
REFERENCE/DOCKET NUMBER: 07254/044W01
TELECOMMUNICATION INFORMATION:
TELEPHONE: 619/678-5070
TELEFAX: 619/678-5099
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 650 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-800-291B-4

```

Query Match 4.7%; Score 89; DB 4; Length 650;  
Best Local Similarity 21.9%; Pred. No. 0.2;  
Matches 62: Conservative 49; Mismatches 102: Indels

[illegible]

```
QY 232 SSV-----CSLLYFGDERYQKGLS-----LWPIVFISSG 259
: : : : : : : : : : : : : : : : : : : : : : : : : :
: : : : : : : : : : : : : : : : : : : : : : : : : :
Db 435 AFLAVLDFINAALSWMIGDMVDIQGLSFQLCSYILRPPVAFLMG 477
: : : : : : : : : : : : : : : : : : : : : : : : : :
: : : : : : : : : : : : : : : : : : : : : : : : : :
RESULT 10
US-09-256-703-2
; Sequence 2, Application US/09256703
; Patent No. 6294379
; GENERAL INFORMATION:
; APPLICANT: Dong, Jian-yun
; APPLICANT: Kan, Yuet Wai
; APPLICANT: The Regents of the University of California
; TITLE OF INVENTION: Efficient AAV Vectors
; FILE REFERENCE: 023070-0849100S
; CURRENT APPLICATION NUMBER: US/09/256.703
; CURRENT FILING DATE: 1999-02-24
; PRIOR APPLICATION NUMBER: US 60/075,980
; PRIOR FILING DATE: 1998-02-25
; NUMBER OF SEQ ID NOS: 7
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2
; LENGTH: 1476
; TYPE: PRT
; ORGANISM: Homo sapiens
; OTHER INFORMATION: truncated cystic fibrosis transmembrane
US-09-256-703-2

Query Match 4.6%; Score 87; DB 4; Length 1476;
Best Local Similarity 20.8%; Pred. No. 1,2;
Matches 83; Conservative 63; Mismatches 112; Indels 142; Gaps 24;

QY 24 ADMVSCVGMFFVLGMFECTAEMSVTLTLOHGVVPAEGLPSGSRILYHYGV-----KD 78
: : : : : : : : : : : : : : : : : : : : : : : : : :
Db 923 ADTLANGFFRGLPLVHTLITVSKILHKKMLHSLV---QAPMSTLTKAGGILNRFSD 979
: : : : : : : : : : : : : : : : : : : : : : : : : :
QY 79 LA-----TVF-FYMLVAIIH-----ATIQEYVLDK-----LSRRLQLTGKGQ 115
: : : : : : : : : : : : : : : : : : : : : : : : : :
Db 980 IAILDDLPLTFIDTQILLVIGAIIVAVVLPYFIVATVPVIVAFIMLRAYFLOTSSQ 1039
: : : : : : : : : : : : : : : : : : : : : : : : : :
QY 116 NK-LNEAGQLSVF-YIVSGIWGMILASENCLSDPTLLWKSQPHNMFTQMKFFYISOL- 172
: : : : : : : : : : : : : : : : : : : : : : : : : :
Db 1040 LKOLESEGRSPIFTHLTVSLKGLWTLRA-----FGRQPY-----FETLFHKA 1085
: : : : : : : : : : : : : : : : : : : : : : : : : :
QY 173 -AYWFSHPPELYFQKVRKQDIPGOLIIYGLHFGHGGAYLLYL-----NHLGLLLM-- 223
: : : : : : : : : : : : : : : : : : : : : : : : : :
Db 1086 TANWF-----LYLSTLRWFQRIEMIFV---IFFIATVFISILTTGEGEGRVGIILTAM 1137
: : : : : : : : : : : : : : : : : : : : : : : : : :
QY 224 -----LHYAV-----ELLSSVCSLLYFGD-----ERYQKGL----- 249
: : : : : : : : : : : : : : : : : : : : : : : : : :
Db 1138 NIMSTLOWAVNSSIDVDSLMSRSVRVFKFIDMPTGKPTKTKPKYKNGQLSKVMIENSH 1197
: : : : : : : : : : : : : : : : : : : : : : : : : :
QY 250 ----SLWPIVFTSGRLVTL--IVSVTVTVGLHLAGTNRNGNAL----- 285
: : : : : : : : : : : : : : : : : : : : : : : : : :
Db 1198 VKKDDIWP-----SGQMTVKDLTKAYTEG-----GNAILNISFISPGORVCLLG 1244
: : : : : : : : : : : : : : : : : : : : : : : : : :
QY 286 ----SGNVNVLAAKIAVLSSSCSIQV-YITWTLTTVWLQRW 321
: : : : : : : : : : : : : : : : : : : : : : : : : :
Db 1245 RTGSGKSTLLSAFLRLNTEGEIQIDGVSW--DSITLQOW 1282
: : : : : : : : : : : : : : : : : : : : : : : : : :
RESULT 11
US-08-951-912-4
; Sequence 4, Application US/08951912
; Patent No. 5972995
; GENERAL INFORMATION:
; APPLICANT: Fischer, Horst
; APPLICANT: Illek, Beate
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR CYSTIC
; NUMBER OF SEQUENCES: 6

CORRESPONDENCE ADDRESS:
ADDRESSEE: SEED and BERRY LLP
STREET: 6300 Columbia Center, 701 Fifth Avenue
CITY: Seattle
STATE: Washington
COUNTRY: USA
ZIP: 98104
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/951,912
FILING DATE: 16-OCT-1997
CLASSIFICATION: 514
ATTORNEY/AGENT INFORMATION:
NAME: Maki, David J.
REGISTRATION NUMBER: 31,392
REFERENCE/DOCKET NUMBER: 200116.403
TELECOMMUNICATION INFORMATION:
TELEPHONE: (206) 622-4900
TELEFAX: (206) 682-6031
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 1479 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-951-912-4

Query Match 4.6%; Score 87; DB 2; Length 1479;
Best Local Similarity 20.8%; Pred. No. 1,3;
Matches 83; Conservative 63; Mismatches 112; Indels 142; Gaps 24;

QY 24 ADMVSCVGMFFVLGMFECTAEMSVTLTLOHGVVPAEGLPSGSRILYHYGV-----KD 78
: : : : : : : : : : : : : : : : : : : : : : : : : :
Db 922 ADTLANGFFRGLPLVHTLITVSKILHKKMLHSLV---QAPMSTLTKAGGILNRFSD 978
: : : : : : : : : : : : : : : : : : : : : : : : : :
QY 79 LA-----TVF-FYMLVAIIH-----ATIQEYVLDK-----LSRRLQLTGKGQ 115
: : : : : : : : : : : : : : : : : : : : : : : : : :
Db 979 IAILDDLPLTFIDTQILLVIGAIIVAVVLPYFIVATVPVIVAFIMLRAYFLOTSSQ 1038
: : : : : : : : : : : : : : : : : : : : : : : : : :
QY 116 NK-LNEAGQLSVF-YIVSGIWGMILASENCLSDPTLLWKSQPHNMFTQMKFFYISOL- 172
: : : : : : : : : : : : : : : : : : : : : : : : : :
Db 1039 LKOLESEGRSPIFTHLTVSLKGLWTLRA-----FGRQPY-----FETLFHKA 1084
: : : : : : : : : : : : : : : : : : : : : : : : : :
QY 173 -AYWFSHPPELYFQKVRKQDIPGOLIIYGLHFGHGGAYLLYL-----NHLGLLLM-- 223
: : : : : : : : : : : : : : : : : : : : : : : : : :
Db 1085 TANWF-----LYLSTLRWFQRIEMIFV---IFFIATVFISILTTGEGEGRVGIILTAM 1136
: : : : : : : : : : : : : : : : : : : : : : : : : :
QY 224 -----LHYAV-----ELLSSVCSLLYFGD-----ERYQKGL----- 249
: : : : : : : : : : : : : : : : : : : : : : : : : :
Db 1137 NIMSTLOWAVNSSIDVDSLMSRSVRVFKFIDMPTGKPTKTKPKYKNGQLSKVMIENSH 1196
: : : : : : : : : : : : : : : : : : : : : : : : : :
QY 250 ----SLWPIVFTSGRLVTL--IVSVTVTVGLHLAGTNRNGNAL----- 285
: : : : : : : : : : : : : : : : : : : : : : : : : :
Db 1197 VKKDDIWP-----SGQMTVKDLTKAYTEG-----GNAILNISFISPGORVCLLG 1243
: : : : : : : : : : : : : : : : : : : : : : : : : :
QY 286 ----SGNVNVLAAKIAVLSSSCSIQV-YITWTLTTVWLQRW 321
: : : : : : : : : : : : : : : : : : : : : : : : : :
Db 1244 RTGSGKSTLLSAFLRLNTEGEIQIDGVSW--DSITLQOW 1281
: : : : : : : : : : : : : : : : : : : : : : : : : :

RESULT 12
US-09-174-077-4
; Sequence 4, Application US/09174077
; Patent No. 6329422
; GENERAL INFORMATION:
; APPLICANT: Fischer, Horst
; APPLICANT: Illek, Beate
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR CYSTIC FIBROSIS THERAPY
```

FILE REFERENCE: 200116.403C1  
 CURRENT APPLICATION NUMBER: US/09/174,077  
 CURRENT FILING DATE: 1998-10-16  
 EARLIER APPLICATION NUMBER: US 08/951,912  
 EARLIER FILING DATE: 1997-10-16  
 NUMBER OF SEQ ID NOS: 6  
 SOFTWARE: Patent In Ver. 2.0  
 SEQ ID NO 4  
 LENGTH: 1479  
 TYPE: PRT  
 ORGANISM: Homo sapiens  
 US-09-174-077-4

Query Match 4.6%; Score 87; DB 4; Length 1479;  
 Best Local Similarity 20.8%; Pred. No. 1.3;  
 Matches 83; Conservative 63; Mismatches 112; Indels 142; Gaps 24;

QY 24 ADMVSCVGMFFVLGLMFECTAEMSVFLTLQGVVPAEGLPSGSRITLYHYGV-----KD 78  
 DB 922 ADTLAMGFRGLPLVHTLITVSKILHHKMLHSVL---QAPMSTLNTLKAGGILNRFSKD 978  
 QY 79 LA-----TVF-FYMLVAIIH-----ATIOEYVLDK-----LSRRLQLTGKQ 115  
 DB 979 IAILDDLLPTIFDFIQLLLIVIGAIYVAVLQPIYFVATVPVIVAFIMLRAYFLQTSQ 1038  
 QY 116 NK-LNPAQOLSVF-YIVSGIMGMILASENCSDPTLWKSPHNMFTQMKFFYISOL- 172  
 DB 1039 LKOLESEGRSPITLVTSLKGLTLRA-----FGRQPY-----FETLFKALNLH 1084  
 QY 173 -AYWHSFPPELYFQVKRKODIPGOLYIYGLHFIHGAYLLYL-----NHLGLLLLM-- 223  
 DB 1085 TANWF-----LYLSTLWFMQRIEMIFV---IFFTAVTIFISLTTGEGEGRVGIILTLAM 1136  
 QY 224 -----LHYAV-----ELLSSVCSLLYFGD-----ERYQKGL----- 249  
 DB 1137 NIMSTLQWAVNSSIDVSLMRSVSRVFKFIDMPTGKPTKTKPKYKNGQSKVMIIENSH 1196  
 QY 250 ----SLWPIVIFISGRVLT--IVSVTVGLHLAGTRNGNAL----- 285  
 DB 1197 VKDDIWP-----SGGQMTVKDTAKYTEG-----GNAILNISFSISPGORVGLLG 1243  
 QY 286 ---SGNVNVAIAKIAVLSSCSIQV-YITWTLTVMQLRW 321  
 DB 1244 RTGSGKSTLSAFLRLNTEGEIQDGVSW--DSITLQW 1281

RESULT 13  
 US-07-637-621-2  
 Sequence 2, Application US/07637621  
 Patent No. 5407796  
 GENERAL INFORMATION:  
 APPLICANT: cutting, gary  
 APPLICANT: antonarakis, stylanos e  
 APPLICANT: kazazian jr., haig h  
 TITLE OF INVENTION: CYSTIC FIBROSIS MUTATION CLUSTER  
 NUMBER OF SEQUENCES: 4  
 CORRESPONDENCE ADDRESS:  
 ADDRESSEE: Banner, Birch, McKie and Beckett  
 STREET: 1001 G Street, N.W.  
 CITY: Washington, D.C.  
 COUNTRY: USA  
 ZIP: 20001  
 COMPUTER READABLE FORM:  
 MEDIUM TYPE: Floppy disk  
 COMPUTER: IBM PC compatible  
 OPERATING SYSTEM: PC-DOS/MS-DOS  
 SOFTWARE: Patent In Release #1.0, Version #1.25  
 CURRENT APPLICATION DATA:  
 APPLICATION NUMBER: US/07/637,621  
 FILING DATE: 19910104  
 CLASSIFICATION: 435  
 ATTORNEY/AGENT INFORMATION:

NAME: kagan, sarah a  
 REGISTRATION NUMBER: 32,141  
 REFERENCE/DOCKET NUMBER: 1107.030010  
 TELECOMMUNICATION INFORMATION:  
 TELEPHONE: 202-508-9100  
 TELEFAX: 202-508-9100  
 INFORMATION FOR SEQ ID NO: 2:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 1480 amino acids  
 TYPE: AMINO ACID  
 TOPOLOGY: linear  
 MOLECULE TYPE: protein  
 ORIGINAL SOURCE:  
 ORGANISM: HOMO SAPIENS  
 US-07-637-621-2

Query Match 4.6%; Score 87; DB 1; Length 1480;  
 Best Local Similarity 20.8%; Pred. No. 1.3;  
 Matches 83; Conservative 63; Mismatches 112; Indels 142; Gaps 24;

QY 24 ADMVSCVGMFFVLGLMFECTAEMSVFLTLQGVVPAEGLPSGSRITLYHYGV-----KD 78  
 DB 923 ADTLAMGFRGLPLVHTLITVSKILHHKMLHSVL---QAPMSTLNTLKAGGILNRFSKD 979  
 QY 79 LA-----TVF-FYMLVAIIH-----ATIOEYVLDK-----LSRRLQLTGKQ 115  
 DB 980 IAILDDLLPTIFDFIQLLLIVIGAIYVAVLQPIYFVATVPVIVAFIMLRAYFLQTSQ 1039  
 QY 116 NK-LNPAQOLSVF-YIVSGIMGMILASENCSDPTLWKSPHNMFTQMKFFYISOL- 172  
 DB 1040 LKOLESEGRSPITLVTSLKGLTLRA-----FGRQPY-----FETLFKALNLH 1085  
 QY 173 -AYWHSFPPELYFQVKRKODIPGOLYIYGLHFIHGAYLLYL-----NHLGLLLLM-- 223  
 DB 1086 TANWF-----LYLSTLWFMQRIEMIFV---IFFTAVTIFISLTTGEGEGRVGIILTLAM 1137  
 QY 224 -----LHYAV-----ELLSSVCSLLYFGD-----ERYQKGL----- 249  
 DB 1138 NIMSTLQWAVNSSIDVSLMRSVSRVFKFIDMPTGKPTKTKPKYKNGQSKVMIIENSH 1197  
 QY 250 ----SLWPIVIFISGRVLT--IVSVTVGLHLAGTRNGNAL----- 285  
 DB 1198 VKDDIWP-----SGGQMTVKDTAKYTEG-----GNAILNISFSISPGORVGLLG 1244  
 QY 286 ---SGNVNVAIAKIAVLSSCSIQV-YITWTLTVMQLRW 321  
 DB 1245 RTGSGKSTLSAFLRLNTEGEIQDGVSW--DSITLQW 1282

RESULT 14  
 US-08-136-742A-2  
 Sequence 2, Application US/08136742A  
 Patent No. 5670488  
 GENERAL INFORMATION:  
 APPLICANT: Gregory, R.J., Armentano, D., Couture, L.A., Smith,  
 APPLICANT: A.E.  
 TITLE OF INVENTION: GENE THERAPY FOR CYSTIC FIBROSIS  
 NUMBER OF SEQUENCES: 10  
 CORRESPONDENCE ADDRESS:  
 ADDRESSEE: BRUMBAUGH, GRAVES, DONOHUE & RAYMOND  
 STREET: 30 ROCKEFELLER PLAZA  
 CITY: NEW YORK  
 STATE: NEW YORK  
 COUNTRY: USA  
 ZIP: 10112  
 COMPUTER READABLE FORM:  
 MEDIUM TYPE: Floppy disk  
 COMPUTER: IBM PC compatible  
 OPERATING SYSTEM: PC-DOS/MS-DOS  
 SOFTWARE: ASCII  
 CURRENT APPLICATION DATA:  
 APPLICATION NUMBER: US/08/136,742A

;; FILING DATE: 02-DEC-1993  
;; CLASSIFICATION: 514  
;; PRIOR APPLICATION DATA:  
;; APPLICATION NUMBER: US 07/985,478  
;; FILING DATE: 02-DEC-1992  
;; CLASSIFICATION: 514  
;; ATTORNEY/AGENT INFORMATION:  
;; NAME: Seide, Rochelle K.  
;; REGISTRATION NUMBER: 32,300  
;; REFERENCE/DOCKET NUMBER: A30668 (Genzyme Dkt. IG4-9.11)  
;; TELECOMMUNICATION INFORMATION:  
;; TELEPHONE: (212) 408-2500  
;; TELEFAX: (212) 765-2519  
;; INFORMATION FOR SEQ ID NO: 2:  
;; SEQUENCE CHARACTERISTICS:  
;; TYPE: amino acid  
;; TOPOLOGY: linear  
;; MOLECULE TYPE: protein  
US-08-136-742A-2

Query Match 4.6%; Score 87; DB 1; Length 1480;  
Best Local Similarity 20.8%; Pred. No. 1.3;  
Matches 83; Conservative 63; Mismatches 112; Indels 142; Gaps 24;

QY 24 ADMVSCVGMFVLGLMFEFTAEMSIIVFLTLQHGCVVPAEGLPSGRTLYHYGV-----KD 78  
|| :  
DB 923 ADTLAMGFFRGFLPVHTLTVSKILHKLHLSVL---QAPMSTLNTLKAGGLNRFSD 979  
:  
QY 79 LA-----TVF-FYMLVAIIH-----ATIOEYVLDK-----LSRRLQLTGKQ 115  
:  
DB 980 IAILDDLLPLTIFDFIQLLLIVIGAIWAVLQPIYFVATPVIVAFIMLRAYFLQTSQ 1039  
:  
QY 116 NK-LNEAGQLSVF-YIVSGIWMIIASENCSDPTLLKSKOPHNMTFMKFFIISQL- 172  
|| :  
DB 1040 LKQLESGRSPITFHLVTSKGLWTLRA-----FGRQPY-----FETLFHKALNLH 1085  
:  
QY 173 -AYWFHSPPELYFOKVRKQDIPGQIYIGLHFLHIGGAYLLYL-----NHLGLLLM-- 223  
:  
DB 1086 TANWF-----LYLSTLRWFQMRIEMIFV---IFFIAVTISILTTGEGRGVGIILTAM 1137  
:  
QY 224 -----LHYAV-----ELSSVCSLLYFGD-----ERYOKGL----- 249  
:  
DB 1138 NIMSTLOWAVNSSIDVDSLMSRSVRVFKFIDMPTGKPTKSTKPKYKNGQLSKVMIENSH 1197  
|| :  
QY 250 ----SLMPIVIFISGRVLT--IYVVTVGLHLAGTNRNGAL-----GNAILENISFSISPGQVGLLG 1244  
:  
DB 1245 RTGSGKSTLLSAFLRLNTEGEIQIDGVSW--DSITLQOW 1282

RESULT 15  
US-08-135-809A-2  
; Sequence 2, Application US/08135809A  
; Patent No. 5688677  
; GENERAL INFORMATION:  
; APPLICANT: CHENG, SENG H.  
; APPLICANT: DITULLIO, PAUL  
; APPLICANT: EBERT, KARL M.  
; APPLICANT: MEADE, HARRY M.  
; APPLICANT: SMITH, ALAN E.  
; TITLE OF INVENTION: DEOXYRIBONUCLEIC ACIDS CONTAINING  
; TITLE OF INVENTION: INACTIVATED HORMONE RESPONSIVE ELEMENTS  
; NUMBER OF SEQUENCES: 9  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: GENZYME CORPORATION  
; STREET: ONE MOUNTAIN ROAD  
; CITY: FRAMINGHAM  
; STATE: MASSACHUSETTS

;; COUNTRY: USA  
;; ZIP: 01701  
;; COMPUTER READABLE FORM:  
;; MEDIUM TYPE: Floppy disk  
;; COMPUTER: IBM PC compatible  
;; OPERATING SYSTEM: PC-DOS/MS-DOS  
;; SOFTWARE: PatentIn Release #1.0, Version #1.25  
;; CURRENT APPLICATION DATA:  
;; APPLICATION NUMBER: US/08/135,809A  
;; FILING DATE: 13-OCT-1993  
;; CLASSIFICATION: 800  
;; ATTORNEY/AGENT INFORMATION:  
;; NAME: LASSEN, ELIZABETH  
;; REGISTRATION NUMBER: 31,845  
;; REFERENCE/DOCKET NUMBER: IG4-9.12  
;; TELECOMMUNICATION INFORMATION:  
;; TELEPHONE: (508) 872-8400  
;; TELEFAX: (508) 872-5415  
;; INFORMATION FOR SEQ ID NO: 2:  
;; SEQUENCE CHARACTERISTICS:  
;; TYPE: amino acid  
;; TOPOLOGY: linear  
;; MOLECULE TYPE: protein  
US-08-135-809A-2

Query Match 4.6%; Score 87; DB 1; Length 1480;  
Best Local Similarity 20.8%; Pred. No. 1.3;  
Matches 83; Conservative 63; Mismatches 112; Indels 142; Gaps 24;

QY 24 ADMVSCVGMFVLGLMFEFTAEMSIIVFLTLQHGCVVPAEGLPSGRTLYHYGV-----KD 78  
|| :  
DB 923 ADTLAMGFFRGFLPVHTLTVSKILHKLHLSVL---QAPMSTLNTLKAGGLNRFSD 979  
:  
QY 79 LA-----TVF-FYMLVAIIH-----ATIOEYVLDK-----LSRRLQLTGKQ 115  
:  
DB 980 IAILDDLLPLTIFDFIQLLLIVIGAIWAVLQPIYFVATPVIVAFIMLRAYFLQTSQ 1039  
:  
QY 116 NK-LNEAGQLSVF-YIVSGIWMIIASENCSDPTLLKSKOPHNMTFMKFFIISQL- 172  
|| :  
DB 1040 LKQLESGRSPITFHLVTSKGLWTLRA-----FGRQPY-----FETLFHKALNLH 1085  
:  
QY 173 -AYWFHSPPELYFOKVRKQDIPGQIYIGLHFLHIGGAYLLYL-----NHLGLLLM-- 223  
:  
DB 1086 TANWF-----LYLSTLRWFQMRIEMIFV---IFFIAVTISILTTGEGRGVGIILTAM 1137  
:  
QY 224 -----LHYAV-----ELSSVCSLLYFGD-----ERYOKGL----- 249  
:  
DB 1138 NIMSTLOWAVNSSIDVDSLMSRSVRVFKFIDMPTGKPTKSTKPKYKNGQLSKVMIENSH 1197  
|| :  
QY 250 ----SLMPIVIFISGRVLT--IYVVTVGLHLAGTNRNGAL-----GNAILENISFSISPGQVGLLG 1244  
:  
DB 1198 VKDDIWP-----SGQMTVKDLTKAYTEG-----GNAILENISFSISPGQVGLLG 1244  
QY 286 ---SGNVNLAAKIIVLSSSCSIQV-YITWTLTVWLQW 321  
|| :  
DB 1245 RTGSGKSTLLSAFLRLNTEGEIQIDGVSW--DSITLQOW 1282

Search completed: September 20, 2002, 14:35:11  
Job time: 63 sec

GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 20, 2002, 14:34:08 ; Search time 20.92 Seconds  
(without alignments)  
1667.324 Million cell updates/sec

Title: US-09-807-470-2

Perfect score: 1892  
Sequence: 1 MGLRKKARNPPVLSHFV.....NGVENPNRIDSPPKKKEKAP 363

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 283138 seqs, 96089334 residues

Total number of hits satisfying chosen parameters: 283138

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000  
Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : PIR.71.\*  
1: pir1.\*  
2: pir2.\*  
3: pir3.\*  
4: pir4.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	1150.5	60.8	374	2	S21736	translocating chai
2	1134.5	60.0	374	2	S30034	translocating chai
3	441.5	23.3	371	2	T19417	hypothetical prote
4	434.5	23.0	373	2	T19419	hypothetical prote
5	148.5	7.8	411	2	S46800	LAG1 protein - yea
6	137	7.2	418	2	S30134	hypothetical prote
7	133	7.0	357	2	T40389	longevity assuranc
8	125	6.6	390	2	T38012	longevity-assuranc
9	116.5	6.2	696	2	B86726	hypothetical prote
10	107.5	5.7	614	2	A69845	Na+/H+ antiporter
11	103.5	5.5	393	2	E95954	ABC transporter (n
12	102.5	5.4	522	2	B83987	ABC transporter (p
13	101	5.3	308	2	H86268	hypothetical prote
14	100.5	5.3	360	2	T27324	hypothetical prote
15	100.5	5.3	397	2	D71467	probable tyrosine
16	99.5	5.3	355	2	D81729	Mtr/TnaB/Tyro perm
17	99	5.2	286	2	AG0403	anaerobic dimethyl
18	98.5	5.2	722	2	G83685	hypothetical prote
19	97.5	5.2	397	2	T00098	hypothetical prote
20	97	5.1	372	2	AH0703	probable membrane
21	96.5	5.1	534	2	D71698	cytochrome-c oxida
22	95.5	5.0	333	2	A10050	probable ABC trans
23	95.5	5.0	411	2	G90154	arsenite transport
24	95	5.0	370	2	H64926	probable membrane
25	95	5.0	370	2	C90928	hypothetical prote
26	95	5.0	370	2	G85776	hypothetical prote
27	95	5.0	509	2	G81929	probable iron-upta
28	94.5	5.0	540	1	I49454	sterol O-acyltrans
29	94.5	5.0	547	2	T27253	hypothetical prote

30 94.5 5.0 788 2 S48191 probable ubiquinol  
31 94 5.0 707 2 T09340 hypothetical prote  
32 94 5.0 1224 2 H96615 hypothetical prote  
33 94 5.0 1450 2 JC6139 cystic fibrosis tr  
34 93.5 4.9 519 2 D85437 hypothetical prote  
35 93.5 4.9 521 2 A34331 cytochrome-c oxida  
36 93 4.9 417 2 AB0682 probable membrane  
37 93 4.9 495 2 E81251 NADH dehydrogenase  
38 93 4.9 515 2 B90504 amino acid transpo  
39 93 4.9 531 2 T40575 major facilitator  
40 93 4.9 532 1 E69821 multidrug resistan  
41 93 4.9 542 2 E90604 hypothetical prote  
42 92.5 4.9 353 2 AF0676 hydrogenase-1 Oper  
43 92.5 4.9 723 2 T21869 hypothetical prote  
44 92 4.9 345 2 AC1252 conserved hypotet  
45 92 4.9 447 2 D90012 hypothetical prote

ALIGNMENTS

RESULT 1

S21736  
translocating chain-associating membrane protein - dog  
C:Species: Canis lupus familiaris (dog)  
C:Date: 22-Nov-1993 #sequence\_revision 01-Dec-1995 #text\_change 24-Sep-1999  
C:Accession: S21736  
R:Goerlich, D.; Hartmann, E.; Prehn, S.; Rapoport, T.A.  
A:Title: A protein of the endoplasmic reticulum involved early in polypeptide transla  
A:Reference number: S21736; MUID:92244357  
A:Accession: S21736  
A>Status: preliminary  
A:Molecule type: mRNA  
A:Residues: 1-374 <GOE>  
A:Cross-references: EMBL:X63678; NID:g941; PIDN:CAA45217.1; PID:g942  
C:Superfamily: translocating chain-associating membrane protein  
C:Keywords: transmembrane protein

Query Match 60.8%; Score 1150.5; DB 2; Length 374;  
Best Local Similarity 59.8%; Pred. No. 1.8e-92;  
Matches 223; Conservative 58; Mismatches 81; Indels 11; Gaps 4;

Qy 1 MGLRKKARNPPVLSHFVQNEADVMVSCVGMFFVLGLMFEFTAEMSIVFLTLQHQVWVP 60  
Db 1 MAIRKSTKSPVLSHFVILQNEADIVSCVAMVFLGLMFEITAKASIIIVTLQYNVTL 60  
Qy 61 A-EGLPSGSRRLHYGVKDLATVFFYMLVAIIHATIQEYVLDKLSRRLOLTKGKONKLN 119  
Db 61 ATEQEAQESTSLYYGKDLATVFFYMLVAIIHATIQEYVLDKINRRMHFSKTSKFN 120  
Qy 120 EAGOLSVFVIVSGIWMIIILASCLSDPTLLWKSOPHNMTFMKFFYISOLAYWFHSF 179  
Db 121 ESGGLFVIFSCVGFILISEYISDPILWRAPVPHNLMTFMKFFYIAQLAYWFHAF 180  
Qy 180 PELYFQVRKQDIPGOLYIIGLHLFHIGGAYLLYLNHLGLLLMLHYAVELLSSVCSLLY 239  
Db 181 PELYFQTKKEDIPQLYVIGLYVFIHAGAYLLNHLGLVLLVLYVEVEFLPHISRLFY 240  
Qy 240 FGDERYQKGLSLPIVPIISRLVTLVSVVTVGLHLA-GTNRNGNALSGNVNVAIAKIAV 298  
Db 241 FSDEKYQKGFSLWAVFLVGLRLTLILSVLTGVFGLAARAENQKLDIFSAGFNVLAVRIAV 300  
Qy 299 LSSSCSIQVYITWTLTVVLRQWLEDANLHVCGRKRR-----SRSKGKTGVNGE---NP 349  
Db 301 LASICITQAFMMKFIINFQLRWRREHSTFOAPVVKKKPTVTKGSRKSKGTGVNGVTVIS 360  
Qy 350 NRIDSPPKKKEKA 362  
Db 361 NGADSPNRKKEKS 373

RESULT 2  
S30034  
translocating chain-associating membrane protein - human  
C:Species: Homo sapiens (man)  
C:Date: 06-Jan-1995 #sequence\_revision 06-Jan-1995 #text\_change 24-Sep-1999  
C:Accession: S30034  
R:Goerlich, D.; Hartmann, E.; Prehn, S.; Rapoport, T.A.  
Nature 357, 47-52, 1992  
A:Title: A protein of the endoplasmic reticulum involved early in polypeptide translocation  
A:Reference number: S21736; MUID:92244357  
A:Accession: S30034  
A:Status: preliminary  
A:Molecule type: mRNA  
A:Residues: 1-374 <GOE>  
A:Cross-references: EMBL:XG3679; NID:g37264; PIDN:CAA45218.1; PID:g37265  
C:Superfamily: translocating chain-associating membrane protein

Query Match 60.0%; Score 1134.5; DB 2; Length 374;  
Best Local Similarity 58.7%; Pred. NO. 4.6e-91;  
Matches 219; Conservative 61; Mismatches 82; Indels 11; Gaps 4;

[illegible]

RESULT 3  
T19417  
hypothetical protein C24F3.la - Caenorhabditis elegans  
C:Species: Caenorhabditis elegans  
C:date: 15-Oct-1999 #sequence\_revision 15-Oct-1999 #text\_change 21-Jan-2000  
C:Accession: T19417

submitted to the EMBL Data Library, April 1998  
A:Reference number: Z19122  
A:Accession: T19417  
A:Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: DNA  
A:Residues: 1-371 <WIL>  
A:Cross-references: EMBL:AL022716; PIDN:CAA18770.1; GSPDB:GN00022; CESP:C24F3.1a  
A:Experimental source: clone C24F3  
C:Genetics:  
A:Gene: CESP:C24F3.1a  
A:Map position: 4  
A:Introns: 114/3; 158/2; 363/3  
C:Superfamily: translocating chain-associating membrane protein

Query Match	23.3%	Score	441.5	DB 2	Length	371			
Best Local Similarity	30.0%	Pred. No.	9.8e-31						
Matches	112	Conservative	76	Mismatches	152	Indels	33	Gaps	

  

QY	2	GLRKKNARNPPVL	SHFQVQNHADMV	SCVGMFVLGLM	EGFGEAEMS	IVFLTLQHG	VVVVPA	61
DB	7	GSKSKKKPOPP	ILSHFQVQNHADMV	SCVGMFVLGLM	EGFGEAEMS	IVFLTLQHG	VVVVPA	66
QY	62	EGLPGSRTL	LYHY--GVKDLAT	VEFFYMLNA	IIHATIQBY	VLDKLSRRLQ	LTGKQKNLN	119
DB	67	AVBOGOER	VEVHGYSGL	LDLPAIFFY	SVCMVIVH	AAVVOYEG	LDKTSKST	136
QY	120	EAGQLSVFY	YSGINGM	ILLASENC--	LSDP	TLKWSQP--	HNMMTFOM	175
DB	127	ESGF--OM	FVTYS	TAHAFY	IVSERLDF	SEVKSWL	GYPTEHRVMS	185
QY	176	FSPFPELY	FQKVRKQD	IPGQIY	ITGLH	FIHGAY	LLYLNLHGL	235
DB	186	IHOFFPEY	LQKLKRDE	IRQKRSQ	AILHIA	FIATIS	AYFFNTRV	245
QY	236	SLLYGDERY	QKGLS----	LWPIVF	ISGR	LVTLV	SVSVT--	288
DB	246	RFAHEVG--	--RGLSD	PAFKLFNG	SFVYL	VRGSLII	AVNTFWY	301
QY	289	VNVYLA	AKTAVL	SSCSIQ	YITWT	TLTWTW	LQWLED	348
DB	302	FNTAVIR	LNVLLAV	NVYLQ	LFLLYS	FVV-----	FHM--GR	349
QY	349	PNRIDSP	PKKKEK	361				
DB	350	AAAV---	--PKKEK	359				

## RESULTS

RESULT 4

T19419  
hypothetical protein C24F3.1b - *Caenorhabditis elegans*  
C:Species: *Caenorhabditis elegans*  
C:Date: 15-Oct-1999 #sequence\_revision 15-Oct-1999 #text\_change 21-Jan-2000  
C:Accession: T19419  
R:McMurray, A.  
submitted to the EMBL Data Library, April 1998  
A:Reference number: Z19122  
A:Accession: T19419  
A:Status: preliminary; translated from GB/EMBL/DDBJ  
A:Molecule type: DNA  
A:Residues: 1-373 <WIL>  
A:Cross-references: EMBL:AL022716; PIDN:CAAL8772.1; GSPDB:GN00022; CESP:C24F3.1b  
A:Experimental source: clone C24F3  
C:Genetics:  
A:Gene: CESP:C24F3.1b  
A:Map position: 4  
A:Introns: 114/3; 158/2; 365/3  
C:Superfamily: translocating chain-associating membrane protein

Query Match.	23.08;	Score	434.5;	DB	2;	Length	373;
Best Local Similarity	30.1%;	Pred. No.	4e-30;				
Matches 113; Conservative	75;	Mismatches	151;	Indels	37;	Gaps	12;
Qy	2	GLRKKNARNPPYLSHEFWQNHADWVSCVGMFFVGLGMFEETGAEMSVIFLTLQHGVVVPA	61				
Db	7	GSKAKKKPQPTLSHEFTIQRGDIOMSCVMVFVGLGMFPLTHLSLSLETAIQAQNGTIV	66				
Qy	62	EGLPSGSRTLYHW--GVKDLATVFVYMLVAIIITHATIQEVVLDKLSRRQLTKGQNKLN	119				
Db	67	AVEQGERVHGVLGSLDLDPAIFYSVCWIVHAVVQEGLDKISKTKLSKVSTFRG	126				
Qy	120	EAGLSVFIYVINGIMTILASENCLSD-----PTLWKSQP--HNNMTQMKFFYISQL	172				
Db	127	ESPH-QMFFTYTYSIAHAFYIVYSER-LEDFSEVKSFSVWLGYPTTEHRVMSAAVYKLAFFQI	184				
Qy	173	AVWFHSFPELPQKYRKQDIPGQLIYITGLHFIHGAYLLYLNHLGLLMLMLHYAVELLS	232				

Db 185 SYWHPPEFYQLKRDREINQSVQALHIAFIATFNFTRVGLAITLEYITQLIF 244  
Qy 233 SVCSLLYFGDERYQKGLS-----LWPIVFISGRVLTIVSVWT--VGLHLAGNRNGNAL 285  
Db 245 HIARFAHFGV---RKGLSDPAFLFNGSVLVRGLSGIIIAVMTFWYGLRQA-ESPVDIS 300  
Qy 286 SGNVNVLAAKTAVLSSSCSIOVYITWTTLTVLQRMLEDANLHVCGRKRSRKRGTENG 345  
Db 301 AGNFNTAVIRLVNLAVVLLQLQLFLLYSFVV-----FHM-GRFRESNAKKEKKKS 348  
Qy 346 VENPNRIDSPPKKKEK 361  
Db 349 AAAAAAV---PKKEKK 361

## RESULT 5

S46800  
N:LAG1 protein - yeast (Saccharomyces cerevisiae)  
N:Alternate names: hypothetical protein YHL003c  
C:Species: Saccharomyces cerevisiae  
C:Date: 28-Oct-1994 #sequence\_revision 28-Oct-1994 #text\_change 23-Mar-2001  
C:Accession: S46800; A54012  
R:Favellio, T.  
A:Submitted to the EMBL Data Library, June 1994  
A:Description: The sequence of S. cerevisiae cosmid 9780.  
A:Reference number: S46797  
A:Accession: S46800  
A:Molecule type: DNA  
A:Residues: 1-411 <FAV>  
A:Cross-references: EMBL:U10555; NID:g500813; PIDN:AAB68429.1; PID:g500820; MIPS:YHL003c  
R:D'mello, N.P.; Childress, A.M.; Franklin, D.S.; Kale, S.P.; Pinswasdi, C.; Jazwinski, J. Biol. Chem. 269, 15451-15459, 1994  
A:Title: Cloning and characterization of LAG1, a longevity-assurance gene in yeast.  
A:Reference number: A54012; MUID:94253121  
A:Accession: A54012  
A:Molecule type: DNA  
A:Residues: 1-172, 'IV', 175-219, 'C', 221-300, 'TEISGI', 314, 'EKQE', 315, 'DSNDNPTE', 324, 'A', 324  
'A', 381, 'AQQR', 386, 'L', 388, 'NRLARNER' <DAM>  
A:Cross-references: GB:U08133  
C:Genetics:  
A:Gene: SGD:LAG1  
A:Cross-references: SGD:S0000995; MIPS:YHL003c  
A:Map position: 8L  
C:Function:  
A:Description: involved in dextermiation of longevity  
C:Superfamily: hypothetical protein YKL008c  
C:Keywords: transmembrane protein

Query Match 7.8%; Score 148.5; DB 2; Length 411;  
Best Local Similarity 22.0%; Pred. No. 3.6e-05;  
Matches 68; Conservative 46; Mismatches 106; Indels 89; Gaps 11;

Qy 26 MVSQGVFFVLGLMFEGETAEMSVFLTQGVVVPVAGLPSGRTLYHYGVKDLATVPFY 85  
Db 91 LVCVSYAFSLSGNTESNPLHMFVAISYQ-----VDGTD-----YAKGDKLSVFPFY 139  
Qy 86 MLVAIIHATIQEYVLDKLRRLQKGNLNEAGQLSVFYI-VSGIWMIL-ASEN 143  
Db 140 MIFTFLEFLMDVVIREFVYLVNTSEHQKRMLEQMTAIFYCGVSGPGLYIMVHSD 199  
Qy 144 CLSDPTLLWKSQPHMMTFQMKFFYSIQLAY-----FHSFPE 181  
Db 200 WLFKTKPMYRYPVITNPFELFKIFYLGOAAFAQAQCVLQLKPKRKYKELVPHHIVT 259  
Qy 182 L-----YFQVRKQDIPGQLIYI-----GLAHLEH 205  
Db 260 LLLIWSVVFHTKM-----GLAIYITMDVSDFFLSLSTNLNSVTFPFVGLFVF- 312  
Qy 206 IGGAYLLYLNHLGLLLMLHYAVELSSVCSGLLYFGDERYQKGLSLWPIVFISGRVLT 265  
Db 313 ----FWIYLRHVIRILMSVLEFRHEGNYVLNFATQYKCHISL-PIVFV-----LI 361

Qy 266 VSVVTVGLH 274  
Db 362 AALQLVNLV 370  
RESULT 6  
S30134  
N:hypothetical protein YKL008c - yeast (Saccharomyces cerevisiae)  
N:Alternate names: hypothetical protein YKL156  
C:Species: Saccharomyces cerevisiae  
C:Date: 18-Jun-1993 #sequence\_revision 18-Jun-1993 #text\_change 24-Sep-1999  
C:Accession: S30134; S37819; SI7017  
R:Boyer, J.; Pascolo, S.; Richard, G.F.; Dujon, B.  
A:Yeast 9, 279-287, 1993  
A:Title: Sequence of a 7.8 kb segment on the left arm of yeast chromosome XI reveals mammalian UOG-1 gene.  
A:Reference number: S30132; MUID:93255906  
A:Accession: S30134  
A:Molecule type: DNA  
A:Residues: 1-418 <BOY>  
A:Cross-references: GB:S59773; NID:g300231; PIDN:AAC60549.1; PID:g300232  
A:Experimental source: strain S288C  
R:Boyer, J.; Pascolo, S.; Richard, G.F.; Ghazvini, M.; Collea, L.; Thierry, A.; Moi submitted to the Protein Sequence Database, March 1994  
A:Reference number: S37813  
A:Accession: S37819  
A:Molecule type: DNA  
A:Residues: 1-418 <BO2>  
A:Cross-references: EMBL:228008; NID:g485982; PIDN:CAA81843.1; PID:g485983; MIPS:YKL1  
A:Experimental source: strain S288C  
R:Amatruda, J.F.; Gattermeier, D.G.; Cooper, J.A.  
A:Submitted to the EMBL Data Library, August 1991  
A:Description: Yeast capping protein.  
A:Reference number: SI7016  
A:Accession: SI7017  
A:Molecule type: DNA  
A:Residues: 1-149 <AMA>  
A:Cross-references: EMBL:X61398; NID:g455515; PIDN:CAA43670.1; PID:g3445  
C:Genetics:  
A:Map position: 11L  
C:Superfamily: hypothetical protein YKL008c  
C:Keywords: transmembrane protein

Query Match 7.2%; Score 137; DB 2; Length 418;  
Best Local Similarity 23.5%; Pred. No. 0.00037;  
Matches 61; Conservative 44; Mismatches 93; Indels 62; Gaps 11;

Qy 67 GSRTLYHYGVKDLATVFFYMLVAIIHATIQEYVLDKLRRLQKGNLNEAGQLSV 126  
Db 121 GDTNAYKGINDLCFVYFMYIFFELREFLMDVVIREFLHVTPFAIRLHVTSKHKRIKIMEQYAI 180  
Qy 127 FYI-VSGIWMIL-ASENCLSDPTLLWKSQPHMMTFQMKFFYSIQLAY----- 175  
Db 181 FYTCVSGPFGIYCMYHSDLMFFNTKAMRYTPDFTNPELFKFVYLGAAFAQAQCVL 240  
Qy 176 -----FHSFPEL-----YFQVRKQDIPGQLIYIIGLHFGIGAYLLYLNH 216  
Db 241 QLEKPRKDHNLAFHHIVTLLIWSVVFHTKMGLP---IYITMDVSDFLSFKTLNV 297  
Qy 217 L--GLL-----LLMLHYA-VELLSVCS-----LLYFGDERYQKGLSLWPI 254  
Db 298 LDSGLAFPSFAIFVAVIYLRHYINLKLMSVLTQFTGNTGNYVLNFATQYKCHISL-PI 356  
Qy 255 VFISGRVLTIVSVVTVGLH 274  
Db 357 VFFV-----LIGALQLVNLV 370

## RESULT 7

T40389  
N:longevity assurance protein homolog SPBC3E7.15c - fission yeast (Schizosaccharomyces  
C:Species: Schizosaccharomyces pombe



C>Date: 03-Dec-1999 #sequence\_revision 03-Dec-1999 #text\_change 31-Jan-2000  
 C:Accession: T40389; T40499  
 R:Lyne, M.; Wood, V.; Rajandream, M.A.; Barrell, B.G.; Brown, D.; Churcher, C.M.  
 submitted to the EMBL Data Library, May 1998  
 A:Reference number: Z21924  
 A:Accession: T40389  
 A:Status: preliminary; translated from GB/EMBL/DDBJ  
 A:Molecule type: DNA  
 A:Residues: 1-357 <LYN>  
 A:Cross-references: EMBL:AL023534; PIDN:CAA19018.1; GSPDB:GN00067; SPDB:SPBC3E7.15c  
 A:Experimental source: strain 972h-; cosmid c3E7  
 R:Gwilliam, R.; Rajandream, M.A.; Barrell, B.G.; Skelton, J.; Churcher, C.M.  
 submitted to the EMBL Data Library, September 1998  
 A:Reference number: Z21933  
 A:Accession: T40499  
 A:Status: preliminary; translated from GB/EMBL/DDBJ  
 A:Molecule type: DNA  
 A:Residues: 1-357 <GWT>  
 A:Cross-references: EMBL:AL031534; PIDN:CAA20722.1; GSPDB:GN00067; SPDB:SPBC4F6.02c  
 A:Experimental source: strain 972h-; cosmid c4F6  
 C:Genetics:  
 A:Gene: SPBC3E7.15c; SPBC4F6.02c  
 A:Map position: 2

Query Match 7.0%; Score 133; DB 2; Length 357;  
 Best Local Similarity 20.5%; Pred. No. 0.00068;  
 Matches 48; Conservative 45; Mismatches 117; Indels 24; Gaps 7;  
 QY 41 EGTAEISVFTL-----HQVVVPAEGLPSGRTLYHYGVKDLATVFFYML 87  
 DB 29 ERTWIVPLILLTLLGVFVNPNGYIKYIPL-SYPIPGTNPAGYKGRDLIAFLFYAL 87  
 QY 88 VALIIHATIQEVLDKLSRRQL-TKGKQNLNEAGQLSVFYIVSGIWMIL-ASENCL 145  
 DB 88 FTTCREFIMQITIRGRHNIRAPKLRFEEQATCTLYFTVMGSLYVMKQTPMWF 147  
 QY 146 SDPTLLKWSQPHNMFTQMKFFYISQLAYWPHSPPELYQ--KVRKODIPQIYIIGLHL 203  
 DB 148 FNTDAFWEYEPHYHVGSKAFYIEAYWIOQALVLILQLEKPK-DFRELVVHIIITL 206  
 QY 204 FHIGAYLLYLNHLGLLLMLHYAVELLSSVCSLLYFGDERYKQGLSLWPIVI 257  
 DB 207 LLIGLSYFHTWIGLAVITMDTSIWLAKSLC-----NYNTVIVYPIFVI 255

RESULT 8  
 T38012  
 longevity-assurance protein 1 - fission yeast (Schizosaccharomyces pombe)  
 C:Species: Schizosaccharomyces pombe  
 C>Date: 03-Dec-1999 #sequence\_revision 03-Dec-1999 #text\_change 03-Dec-1999  
 C:Accession: T38012  
 R:Churcher, C.M.  
 submitted to the EMBL Data Library, September 1997  
 A:Reference number: Z21761  
 A:Accession: T38012  
 A:Status: preliminary; translated from GB/EMBL/DDBJ  
 A:Molecule type: DNA  
 A:Residues: 1-390 <CHU>  
 A:Cross-references: EMBL:Z99258; PIDN:CAB16359.1; GSPDB:GN00066; SPDB:SPAC1A6.09c  
 A:Experimental source: strain 972h-; cosmid c1A6  
 C:Genetics:  
 A:Gene: lag1; SPDB:SPAC1A6.09c  
 A:Map position: 1

Query Match 6.6%; Score 125; DB 2; Length 390;  
 Best Local Similarity 23.9%; Pred. No. 0.0037;  
 Matches 57; Conservative 35; Mismatches 100; Indels 46; Gaps 11;  
 QY 65 PGSGRTLYHYGVKDLATVFFYMLVAIIHATIQEVLDKLSRRQLTKGKQ-----KLN 119  
 DB 104 PDGS---YKGPCKDACPIFWIVTAFRVIMDYVE----RPFVLNMGVRRNRKVIIRFC 156

QY 120 EAGQLSVFYIVSGIWMITLASENCLSDPTLLKWSQPHNMFTQMKFFYISQLAYWPHSF 179  
 DB 157 EGGYSFFYLLCFWELGLYIVRSSNWSNEKLFEDYQYYMSPLFKAYYLILQLGFWLQOI 216  
 QY 180 PELYFQKVRKOD-----IPQLIYI--GLHLFHIGGAYLLYLNHLGLLL-----ML 224  
 DB 217 IVLHLEQ--RRADHWQMFHAHIVTTCALILSYGNFNLRVGNA--ILYIFDLSDYILSGKML 274  
 QY 225 HYAVELLSSVCSLLYFGDERYKQGLSLWPIVFIISGRITVIVSVTV-----GLHL 275  
 DB 275 KYLG--FGKICDYLF-----GIFVSNVYSRHHYLSKILRVVVTNAPEIIGGFHL 322  
 RESULT 9  
 BB6726  
 hypothetical protein yica [imported] - Lactococcus lactis subsp. lactis (strain IL140  
 C:Species: Lactococcus lactis subsp. lactis  
 C>Date: 23-Mar-2001 #sequence\_revision 23-Mar-2001 #text\_change 03-Aug-2001  
 C:Accession: B86726  
 R:Boletín, A.; Wincker, P.; Mauger, S.; Jallón, O.; Malarre, K.; Weissenbach, J.; Eh  
 Genome Res. 11, 731-753, 2001  
 A:Title: The complete genome sequence of the lactic acid bacterium Lactococcus lactis  
 A:Reference number: A86625; MUID:21235186; PMID:11337471  
 A:Accession: B86726  
 A:Status: preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-696 <STO>  
 A:Cross-references: GB:AE005176; PID:gl2723732; PIDN:AAK04908.1; GSPDB:GN00146  
 A:Experimental source: strain IL1403  
 C:Genetics:  
 A:Gene: yica

Query Match 6.2%; Score 116.5; DB 2; Length 696;  
 Best Local Similarity 21.9%; Pred. No. 0.041;  
 Matches 80; Conservative 62; Mismatches 122; Indels 101; Gaps 18;  
 QY 27 VSCVGMFVLGL-MFEGTAEMSIIVFLTLQHGVPVPAEGLPSGRTLYHYGVKDLATVFFY 85  
 DB 21 ILICINFILNRFSGYTGDDFLY----HFVYTG--WPEHLREYHNLWDWLAVHHT 74  
 QY 86 MLVAIIHATIQEVLDKLSRRQLTKGKQNLNEAGQLSVFYIVSGIWMITLASENCL 145  
 DB 75 ML---IWNARMTSIIFEFA--MQIPKFLNIIN-----SLIYVLIGLLINLVSGKKA 124  
 QY 146 SDPTLLKWSQPHNMFTQMKFFYISQLA-----YWFHSFPELYFQKVRKODIPG 194  
 DB 125 LKPS-----HLSLTFLMWFFLPGMGSIVLWVSGATNLYLWFLSLVILFLAFEDIAA 177  
 QY 195 QLIVYIGLHLF-----HIGG--AYLLYL-----NH-----LGILLMLH 225  
 DB 178 RSNWISLGLFILGLTLGLTNEVGATAFLLALLFTIFNRRQPSERVLTQIFGVLAGIG 237  
 QY 226 YAVELLSSV-----CSLLYFGDERYKQGLSLWPIVFIISGRITVIVSVTV----- 267  
 DB 238 FFIQLLLSSGSETQNTYKSGAGFLQHLSDVFTGMQY--SGFLILPILIGGLLYLRIOQ 296  
 QY 268 -----VVTVGLHAGTNRNGNALGNVLAAKI--AVLSSCSIQYVITWTLFTVWL 318  
 DB 297 TEKVTLVITSLLEL-----GSALAGSIALASPAPLWFAFNILLITLILL--I 347  
 QY 319 QRWLE 323  
 DB 348 EAWQE 352

RESULT 10  
 A69845  
 Na+/H+ antiporter homolog yjbQ - Bacillus subtilis  
 C:Species: Bacillus subtilis  
 C>Date: 05-Dec-1997 #sequence\_revision 05-Dec-1997 #text\_change 15-Oct-1999  
 C:Accession: A69845

R;Kunst, F.; Ogasawara, N.; Moszer, I.; Albertini, A.M.; Alloni, G.; Azevedo, V.; Berteis C.; Bron, S.; Brouillett, S.; Bruschini, C.V.; Caldwell, B.; Capuano, V.; Carter, N.M.; Ehrlich, S.D.; Emmerson, P.T.; Entian, K.D.; Errington, J.; Fabret, C.; Ferrari, E. Nature 390, 249-256, 1997

A:Authors: Foulger, D.; Fritz, C.; Fujita, M.; Fujita, Y.; Fuma, S.; Galizzi, A.; Galled iech, J.; Harwood, C.R.; Henaut, A.; Hilbert, H.; Holsappel, S.; Hosono, S.; Hullo, M.F. Koetter, P.; Konigstein, G.; Krogh, S.; Kumano, M.; Kurita, K.; Lapidus, A.; Lardinois, A.; Authors: Lauber, J.; Lazarevic, V.; Lee, S.M.; Levine, A.; Liu, H.; Masuda, S.; Maueel Y, M.; Ogawa, K.; Ogiwara, A.; Oudega, B.; Park, S.H.; Parro, V.; Pohl, T.M.; Portetelle Rieger, M.; Rivolta, C.; Rocha, E.; Roche, B.; Rose, M.; Sadale, Y.; Sato, T.; Scanlon, A.; Authors: Schleich, S.; Schroeter, R.; Scoffone, F.; Sekiguchi, Y.; Sekowska, A.; Seron akeuchi, M.; Tamakoshi, A.; Tanaka, T.; Terpstra, P.; Togononi, A.; Tosato, V.; Uchiyama, T.; Winters, P.; Wipat, A.; Yamamoto, H.; Yamane, K.; Yasumoto, K.; Yata, K.; Yoshida, K A:Authors: Yoshikawa, H.F.; Zumbstein, E.; Yoshikawa, H.; Danchin, A.

A:Title: The complete genome sequence of the Gram-positive bacterium *Bacillus subtilis*. A:Reference number: A69580; MUID:98044033

A:Accession: A69845

A:Status: preliminary; nucleic acid sequence not shown; translation not shown

A:Molecule type: DNA

A:Residues: 1-614 <KUN>

A:Cross-references: GB:299110; GB:AL009126; NID:2633472; PIDN:CAB13021.1; PID:ell83184; A:Experimental source: strain 168

C:Genetics: A:Gene: yjbQ

Query Match 5.7%; Score 107.5; DB 2; Length 614;  
Best Local Similarity 20.5%; Pred. No. 0.21;  
Matches 67; Conservative 58; Mismatches 123; Indels 79; Gaps 14;

Qy 24 ADMVSCVGMF-----FVLGMPEGTAEMSIPTLQHGWWVPA---EGLPS---GS 68  
Db 103 AASVIFVGIFILSLLSYGVFLAGFONAFMTLIISTLSLGVVPTLKEERIMNSNGQ 162  
Qy 69 RLYHYGVKDLATVFF-----YMLVAIIHATQEVLDKLSRRLOLTGK 114  
Db 163 IILLVAVIADLATMILLAVFSSLYGDSGNMMLLILFAAGVVLFFGVRVHRFSVQSM 222  
Qy 115 QNKLEAGQLSVFIVSGINGMI-ILASENCLS---DPTLLMKSQPHNMTFMKFFEYS 170  
Db 223 SKGTQIGTTRAITLIIVIVALSSEGAENILGAFAGVLVSLSPKELVQQLDSF--- 279  
Qy 171 QLAWFFHSFPPELYFQKVRKODIPQOLIYGLHLFHIGGAYLLNHLGLLLMLHLYAVEL 230  
Db 280 --GYGFL-----IPFFVMGVKL-----DIWTLFQDKTILIMIPLLALL 319  
Qy 231 LSSVCSLLYFGDERYQKGLSLW---PIVFISGRLYT---LIVSVTVG--LHLAGTNRN 281  
Db 320 VSKTIIPWY-----LKKWYDNRITFASGFLTSLTSLVIAAATIGQQLHVISTN-- 368  
Qy 282 GNALSGNVNVLAAKIAVLSSSCSIQVY 308  
Db 369 --MSGALILVAVIASITFTPCFKLY 392

RESULT 11  
E96954  
Na/H antiporter (napA) [imported] - Clostridium acetobutylicum  
C:Species: Clostridium acetobutylicum  
C:Date: 14-Sep-2001 #sequence\_revision 14-Sep-2001 #text\_change 30-Sep-2001  
C:Accession: E96954  
R;Nolling, J.; Breton, G.; Omelchenko, M.V.; Markarova, K.S.; Zeng, Q.; Gibson, R.; Lee, J. Bacteriol. 183, 4823-4838, 2001  
A:Title: Genome Sequence and Comparative Analysis of the Solvent-Producing Bacterium *Clostridium acetobutylicum* ATCC824  
A:Reference number: A96900; MUID:21359325; PMID:21359325  
A:Accession: E96954  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-393 <KUR>  
A:Cross-references: GB:AE001437; PIDN:AAK78424.1; PID:gl5023300; GSPDB:GN00168  
A:Experimental source: Clostridium acetobutylicum ATCC824  
C:Genetics:

A:Gene: CAC0444  
C:Superfamily: Aquifex aeolicus Na+/H+-exchanging protein napA1

Query Match 5.5%; Score 103.5; DB 2; Length 393;  
Best Local Similarity 22.4%; Pred. No. 0.28;  
Matches 82; Conservative 59; Mismatches 114; Indels 111; Gaps 21;

Qy 19 MVQNHAD--MVSCVGMFFVLGLMFECTAEMSI-----VFLTLQHGVVVPAEGLPSG 67  
Db 47 LIQANADIKLLSLGVVFF---LMLLAGIETNLDLDELKAGKSSFLIALGAGLIIP---LIVG 100

Qy 68 SRLYHYGVKDLATVFFYMLVAIIHATQEVLDKLSRRLOLTGKONKLEAGOLSVF 127  
Db 101 TLSAYMF---FSNFYENLFFVGVLITAT-----SVSISVQ-TLTGKLRSGINIL 148

Qy 128 --YIVSGINGMIILASENCLSDPTLLMKSQPHNMTFMKFFEYI-----SOLAYWFHSF 179  
Db 149 GAATIDDLGLLITVVLAISSGT---KSHGSSIF-MTFIYIGFICLVSLAIAFLPK 202

Qy 180 P-ELYFQKVRKQDIPQOLIYIGLHLPHIG-----GAYLLYLNHLGLL 220  
Db 203 PIDKLTQKFKPKR-----GLAIFSAALICAFATKLGIAAATGAYI-----CGLV 249

Qy 221 LLMHLY-----AVELLS-SVCSLLYF---GDERYQKGLSLWPIVFISGRLYTLIVSVVT 270  
Db 250 LSPITHKEYIEKRVKIISTSFISPIFFASVGSASVKGILNFVL-----LITLIMFIIA 303

Qy 271 VGLHLAGTNRN-----GNALSGNVNVLAAKIA-----VLSSSCSIQVYIT 310  
Db 304 VICKILGCSASALTFLKFKKSEALQIGVMVSRGEVAITANIGLOAKIIEEIFLPTLIV 363

Qy 311 WTLTV 316  
Db 364 VILTV 369

## RESULT 12

B83987  
ABC transporter (permease) BH2698 [imported] - Bacillus halodurans (strain C-125)  
C:Species: Bacillus halodurans  
C:Date: 01-Dec-2000 #sequence\_revision 01-Dec-2000 #text\_change 15-Jun-2001  
C:Accession: B83987  
R;Takami, H.; Nakasone, K.; Takaki, Y.; Maeno, G.; Sasaki, R.; Fujii, F.; Nucleic Acids Res. 28, 4317-4331, 2000  
A:Title: Complete genome sequence of the alkaliphilic bacterium *Bacillus halodurans* strain C-125  
A:Reference number: A83650; MUID:20512582; PMID:11058132  
A:Accession: B83987  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-522 <STO>  
A:Cross-references: GB:AP001516; GB:BA000004; NID:gi0175192; PIDN:BA06417.1; GSPDB:C-125  
A:Experimental source: strain C-125  
C:Genetics:  
A:Gene: BH2698

Query Match 5.4%; Score 102.5; DB 2; Length 522;  
Best Local Similarity 23.2%; Pred. No. 0.48;  
Matches 67; Conservative 57; Mismatches 66; Indels 99; Gaps 21;

Qy 14 LSHE-FMVQNHADMVSCVGMFFVLGLMFECTAEM-----SIVFLTLQHGVVVPAEGLPS 66  
Db 75 LSHGIFMKQRKEF-----GLFSLGMSYQDMVKMLLENAGIAFLSL-----VVGLLS 123

Qy 67 GS-----RTLYHYGVKDLATVFF---FYMVAIIHATQEVLDKLSRRLOLTGK-- 113  
Db 124 GTVFSRLFFLTSMYYIEVDIDISFQLOCTFY-LYATAFSVI--FIL-AIGKTLFTIRGOK 179

Qy 114'-----KONKLEAGOLSVFIVSGINGMIILASENCLSDPTLLMKSQPHNMTFMKFFEY 168  
Db 180 IIVAKMENRLSEQKQSPW--LGGIGGILVWAS-----LVFLY 215

RESULT 14  
T27324  
hypothetical protein Y6B3B.10 - *Caenorhabditis elegans*  
C:Species: *Caenorhabditis elegans*  
C:Date: 15-Oct-1999 #sequence\_revision 15-Oct-1999 text\_change 15-Oct-1999  
C:Accession: T27324  
R:white, S.  
submitted to the EMBL Data Library, October 1998  
A:Reference number: Z20345  
A:Accession: T27324  
A:Status: preliminary; translated from GB/EMBL/DDBJ  
A:Molecule type: DNA  
A:Residues: 1-360 <WIL>  
A:Cross-references: EMBL:AL032655; PIDN:CAA21723.1; GSPDB:GN00019; CESP:Y6B3B.10  
A:Experimental source: clone Y6B3B  
C:Genetics:  
A:Gene: CESP:Y6B3B.10

A;Gene: **TYFP\_2**  
C;Superfamily: **tyrosine-specific transport protein**

Search completed: September 20, 2002, 14:36:27  
Job time: 139 sec

(

GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: September 20, 2002, 14:35:13 ; Search time 13.4 Seconds  
(without alignments)  
1048.895 Million cell updates/sec

Title: us-09-807-470-2

Perfect score: 1892

Sequence: 1 MGLRKNRNPVLSHEFMV.....NGVENPNRIDSPKKKKAP 363

Scoring table:

BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 105224 seqs, 38719550 residues

Total number of hits satisfying chosen parameters: 105224

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt\_40:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1145.5	60.5	373	1 TRAM_CANFA	Q01685 canis faml
2	1129.5	59.7	373	1 TRAM_HUMAN	Q15629 homo sapien
3	1084.5	57.3	358	1 TRAM_BOVIN	Q09k24 bos taurus
4	754	39.9	370	1 Y557_HUMAN	Q15035 bos sapien
5	148.5	7.8	411	1 LAG1_YEAST	P38703 saccharomyc
6	137	7.2	418	1 YKAB_YEAST	P28496 saccharomyc
7	133	7.0	384	1 YHFE_SCHPO	O59735 schizosacch
8	125	6.6	390	1 LAG1_SCHPO	P78970 schizosacch
9	97.5	5.2	533	1 GLPT_HUMAN	P57057 homo sapien
10	97	5.1	1450	1 CLTR_HUMAN	Q00554 oryctolagus
11	96.5	5.1	534	1 COX1_RICPR	O54069 rickettsia
12	95	5.0	370	1 YDIK_ECOLI	P77175 escherichia
13	94.5	5.0	540	1 SOAL_MOUSE	O61263 mus musculus
14	94.5	5.0	788	1 COXM_SULAC	P39481 sulfolobus
15	94	5.0	272	1 ATP6_BUCAP	O51878 buchnera ap
16	93.5	4.9	521	1 COX1_APILI	P20374 apis mellif
17	93	4.9	532	1 YHCA_BACSU	P54585 bacillus su
18	91.5	4.8	369	1 Y316_MYCPN	O50361 mycoplasma
19	91	4.8	307	1 OGD2_HUMAN	P58182 homo sapien
20	91	4.8	624	1 NKX3_RAT	O9ep00 ratuus norv
21	91	4.8	645	1 NKX3_MOUSE	O99p07 mus musculus
22	90	4.8	513	1 COX1_RABIT	O79429 oryctolagus
23	90	4.8	514	1 COX1_CERS1	O03198 ceratotheri
24	90	4.8	514	1 COX1_EQUAS	P92477 equus asinu
25	90	4.8	514	1 COX1_HORSE	P48659 equus cabal
26	90	4.8	514	1 COX1_RHJUN	O96062 rhinoceros
27	90	4.8	516	1 COX1_GADMO	Q36775 gadus morhu
28	89	4.7	360	1 MRAY_PASMO	P57816 pasteurella
29	89	4.7	430	1 SECY_BACHD	P38375 bacillus ha
30	89	4.7	499	1 MVIN_AQUAE	O67658 aquifex aeo
31	89	4.7	514	1 COX1_BOVIN	P00396 bos taurus
32	89	4.7	514	1 COX1_CANFA	Q92264 canis faml
33	89	4.7	514	1 COX1_FELCA	P48888 felis silve

34	89	4.7	514	1 COX1_SHEEP	O78749 ovis aries
35	89	4.7	644	1 NKX3_HUMAN	Q09c58 homo sapien
36	89	4.7	649	1 CNL1_HUMAN	O00337 homo sapien
37	88.5	4.7	1481	1 CFTR_BOVIN	P35071 bos taurus
38	88.5	4.7	1481	1 CFTR_SHEEP	Q00555 ovis aries
39	88	4.7	433	1 SECY_RICPR	Q92c55 rickettsia
40	88	4.7	515	1 COX1_HIPAM	O92zy9 hippopotamu
41	88	4.7	531	1 YDFG_SCHPO	Q10487 schizosacch
42	87.5	4.6	297	1 CTAAL_BACFI	Q04443 bacillus fi
43	87.5	4.6	515	1 COX1_CHICK	P18943 gallus gall
44	87	4.6	432	1 RFBX_SALTI	O99191 salmonella
45	87	4.6	514	1 COX1_PIG	O79876 sus scrofa

## ALIGNMENTS

### RESULT 1

ID	TRAM_CANFA	STANDARD;	PRT;	373 AA.
AC	Q01685;			
DT	01-JUL-1993 (Rel. 26, Created)			
DT	01-JUL-1993 (Rel. 26, Last sequence update)			
DT	16-OCT-2001 (Rel. 40, Last annotation update)			
DE	TRAM protein (Translocating chain-associating membrane protein).			
OS	Canis familiaris (Dog).			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Carnivora; Fissipedia; Canidae; Canis.			
OX	NCBI_TaxID=9615;			
RN	[1]			
RP	SEQUENCE FROM N.A., AND SEQUENCE OF 1-26 AND 164-184.			
RC	TISSUE=Kidney;			
RX	MEDLINE=92244357; PubMed=1315422;			
RA	Goerlich D., Hartmann E., Prehn S., Rapoport T.A.;			
RT	"A protein of the endoplasmic reticulum involved early in polypeptide translocation."			
RL	Nature 357:47-52(1992).			
CC	FUNCTION: STIMULATORY OR REQUIRED FOR THE TRANSLOCATION OF			
CC	SECRETORY PROTEINS ACROSS THE ER MEMBRANE.			
CC	LOCUS: ENDOPLASMIC RETICULUM MEMBRANE.			
CC	SUBCELLULAR LOCATION: TO THE LASSI FAMILY.			
CC	THIS SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (see http://www.isb-sib.ch/announce/ or send an email to license@isb-sib.ch).			
CC	EMBL; X63678; CAA45217.1; ..			
DR	PIR; S21736; S21736.			
KW	Endoplasmic reticulum; Transmembrane; Glycoprotein; Translocation.			
FT	INIT_MET 0			
FT	DOMAIN 1 28			
FT	TRANSMEM 29 49			
FT	DOMAIN 50 75			
FT	TRANSMEM 76 96			
FT	DOMAIN 97 120			
FT	TRANSMEM 121 141			
FT	DOMAIN 142 158			
FT	TRANSMEM 159 179			
FT	DOMAIN 180 191			
FT	TRANSMEM 192 212			
FT	DOMAIN 213 216			
FT	TRANSMEM 217 237			
FT	DOMAIN 238 250			
FT	TRANSMEM 251 271			
FT	DOMAIN 272 296			
FT	TRANSMEM 297 317			
FT	DOMAIN 318 373			
FT	CARBOHYD 35 55			
SQ	SEQUENCE 373 AA; 43029 MW; ID85808E1D80E835 CRC64;			

Query Match 60.5%; Score 1145.5; DB 1; Length 373;  
Best Local Similarity 59.8%; Pred. No. 2.7e-88;  
Matches 222; Conservative 58; Mismatches 80; Indels 11; Gaps 4;

QY 3 LRKKARNPPVLSHEFWQNHADVSCVGMFFVLGLMFEQTAEMSIIVFLTLQHGVVVPA- 61  
DB 2 IRKSKTSPVLSHEFVQNHADIVSCVAMVFLGLMFEITAKASIIFFVLQYNVLPAT 61  
QY 62 EQLSGSRTLYHYGVKDLATVFFVLMVAIIHATIQBYVLDKLSRLQLTGKQNKLEA 121  
DB 62 EQATESVSLYYIGIKDLATVFFVLMVAIIHATIQBYVLDKLSRLQLTGKQNKLEA 121  
QY 122 GOLSVFYIVSGWGMILASENCUSDPTLLWKQPHNMFTQMKFFYISOLAYWFFHSFPE 181  
DB 122 GOLSAFYLCIWTGFTILISENYISDPTILWRAYPHNLMFTQMKFFYISOLAYWFFHAFFE 181  
QY 182 LYFQKVRKQDIPGOLIVYGLHPLHIGAYLLYNHLGLLLMLHYAVELLSVCSLLYFG 241  
DB 182 LYFQKVRKQDIPGOLIVYGLHPLHIGAYLLYNHLGLLLMLHYAVELLSVCSLLYFG 241  
QY 242 DERYOKGLSLWPIVIFISGRVLTIVSVTVGLHLA-GTNRNGNALSGNVNLAIAVL 300  
DB 242 DEKYOKGFSLWAVLVGLRLTLILSVTVGLHLAENQKLDIFSAGNFVLAIVLA 301  
QY 301 SSCSIQYIITWTLTVWLQRLWEDANLHVCGRKRR-----SRSRKGTEGV- 351  
DB 302 SICITQAFMMWKEINFOLRRRHSHTFOAPVVKKKPTVTKGRSSRKGTEGVNGVTTSNG 361  
QY 352 IDSPPKKKEA 362  
DB 362 ADSPRNKKEKS 372

RESULT 2  
ID TRAM\_HUMAN STANDARD; PRT; 373 AA.  
AC Q15629;  
DT 16-OCT-2001 (Rel. 40, Created)  
DT 16-OCT-2001 (Rel. 40, Last sequence update)  
DT 16-OCT-2001 (Rel. 40, Last annotation update)  
DE TRAM protein (translocating chain-associating membrane protein).  
GN TRAM.  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
OX NCBI\_TaxID=9606;  
RN [1]  
RP MEDLINE=92244357; PubMed=1315422;  
RX Goerlich D., Hartmann E., Prehn S., Rapoport T.A.;  
RA "A protein of the endoplasmic reticulum involved early in polypeptide  
translocation."  
RL Nature 357:47-52(1992).  
RN [2]  
RP SEQUENCE FROM N.A.  
RC Straussberg R.;  
RA TISSUE=Kidney;  
RL Submitted (NOV-2000) to the EMBL/GenBank/DBJ databases.  
CC -1- FUNCTION: STIMULATORY OR REQUIRED FOR THE TRANSLOCATION OF  
SECRETORY PROTEINS ACROSS THE ER MEMBRANE.  
CC -1- SUBCELLULAR LOCATION: Endoplasmic reticulum membrane.  
CC -1- SIMILARITY: BELONGS TO THE LASSI FAMILY.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----

DR EMBL; X63679; CAA45218.1; --  
DR EMBL; BC000687; AAH00687.1; --  
KW MIM; 605190; --  
FT INT\_MET 0 0  
FT DOMAIN 1 28  
FT TRANSMEM 29 49  
FT DOMAIN 50 75  
FT TRANSMEM 76 96  
FT DOMAIN 97 120  
FT TRANSMEM 121 141  
FT DOMAIN 142 158  
FT TRANSMEM 159 179  
FT DOMAIN 180 191  
FT TRANSMEM 192 212  
FT DOMAIN 213 216  
FT TRANSMEM 217 237  
FT DOMAIN 238 250  
FT TRANSMEM 251 271  
FT DOMAIN 272 296  
FT TRANSMEM 297 317  
FT DOMAIN 318 373  
FT CARBOHYD 55 55  
SQ SEQUENCE 373 AA; 42940 MW; C220949AF4EFEDD0 CRC64;

Query Match 59.7%; Score 1129.5; DB 1; Length 373;  
Best Local Similarity 58.8%; Pred. No. 5.8e-87;  
Matches 218; Conservative 61; Mismatches 81; Indels 11; Gaps 4;

QY 3 LRKKARNPPVLSHEFWQNHADVSCVGMFFVLGLMFEQTAEMSIIVFLTLQHGVVVPA- 61  
DB 2 IRKSKTSPVLSHEFVQNHADIVSCVAMVFLGLMFEITAKASIIFFVLQYNVLPAT 61  
QY 62 EQLSGSRTLYHYGVKDLATVFFVLMVAIIHATIQBYVLDKLSRLQLTGKQNKLEA 121  
DB 62 EQATESVSLYYIGIKDLATVFFVLMVAIIHATIQBYVLDKLSRLQLTGKQNKLEA 121  
QY 122 GOLSVFYIVSGWGMILASENCUSDPTLLWKQPHNMFTQMKFFYISOLAYWFFHSFPE 181  
DB 122 GOLSAFYLCVWGTFTILISENYISDPTILWRAYPHNLMFTQMKFFYISOLAYWFFHAFFE 181  
QY 182 LYFQKVRKQDIPGOLIVYGLHPLHIGAYLLYNHLGLLLMLHYAVELLSVCSLLYFG 241  
DB 182 LYFQKVRKQDIPGOLIVYGLHPLHIGAYLLYNHLGLLLMLHYAVELLSVCSLLYFG 241  
QY 242 DERYOKGLSLWPIVIFISGRVLTIVSVTVGLHLA-GTNRNGNALSGNVNLAIAVL 300  
DB 242 NEKYOKGFSLWAVLVGLRLTLILSVTVGLHLAENQKLDIFSAGNFVLAIVLA 301  
QY 301 SSCSIQYIITWTLTVWLQRLWEDANLHVCGRKRR-----SRSRKGTEGV- 351  
DB 302 SICITQAFMMWKEINFOLRRRHSHTFOAPVVKKKPTVTKGRSSRKGTEGVNGVTTSNV 361  
QY 352 IDSPPKKKEA 362  
DB 362 ADSPRNKKEKS 372

RESULT 3  
ID TRAM\_BOVIN STANDARD; PRT; 358 AA.  
AC Q9GKZ4;  
DT 16-OCT-2001 (Rel. 40, Created)  
DT 16-OCT-2001 (Rel. 40, Last sequence update)  
DT 16-OCT-2001 (Rel. 40, Last annotation update)  
DE TRAM protein (translocating chain-associating membrane protein)  
GN TRAM.  
OS Bos taurus (Bovine).  
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;  
OC Bovidae; Bovinae; Bos.

OX NCBI\_TaxID=9913;  
 RN [1]  
 SEQUENCE FROM N.A.  
 RA Clark T.G., Morris J., Akamatsu M., McGraw R.A., Ivarie R.D.;  
 RT "Cloning and sequence analysis of a bovine tram cDNA."  
 RL Submitted (JAN-1995) to the EMBL/GenBank/DBJ databases.  
 CC -!- FUNCTION: STIMULATORY OR REQUIRED FOR THE TRANSLLOCATION OF  
 CC SECRETORY PROTEINS ACROSS THE ER MEMBRANE (BY SIMILARITY).  
 CC -!- SUBCELLULAR LOCATION: ENDOPLASMIC RETICULUM MEMBRANE (BY  
 CC similarity).  
 CC -!- SIMILARITY: BELONGS TO THE LASS1 FAMILY.  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC EMBL: U19578; AAG10391.1; -  
 DR Endoplasmic reticulum; Transmembrane; Glycoprotein; Translocation.  
 KW NON\_TER 1  
 FT DOMAIN <1 13 CYTOPLASMIC (POTENTIAL).  
 FT TRANSMEM 14 34 POTENTIAL.  
 FT DOMAIN 35 60 LUMENAL (POTENTIAL).  
 FT TRANSMEM 61 81 POTENTIAL.  
 FT DOMAIN 82 105 CYTOPLASMIC (POTENTIAL).  
 FT TRANSMEM 106 126 POTENTIAL.  
 FT DOMAIN 127 143 LUMENAL (POTENTIAL).  
 FT TRANSMEM 144 164 CYTOPLASMIC (POTENTIAL).  
 FT DOMAIN 165 176 POTENTIAL.  
 FT TRANSMEM 177 197 LUMENAL (POTENTIAL).  
 FT DOMAIN 198 201 LUMENAL (POTENTIAL).  
 FT TRANSMEM 202 222 POTENTIAL.  
 FT DOMAIN 223 235 CYTOPLASMIC (POTENTIAL).  
 FT TRANSMEM 236 256 POTENTIAL.  
 FT DOMAIN 257 281 LUMENAL (POTENTIAL).  
 FT TRANSMEM 282 302 POTENTIAL.  
 FT DOMAIN 303 358 CYTOPLASMIC (POTENTIAL).  
 FT CARBOHYD 40 40 N-LINKED (GLCNAC... (PROBABLE).  
 SQ SEQUENCE 358 AA; 41403 MW; 25CF9930C4CDDA15 CRC64;  
 Query Match 57.3%; Score 1084.5; DB 1; Length 358;  
 Best Local Similarity 59.9%; Pred. No. 3.1e-83;  
 Matches 214; Conservative 53; Mismatches 79; Indels 11; Gaps 5;  
 QY 17 EPWQNHADVSCVGMFFVLGLMFEPTAEMSIIVFTLQHGVVVPA-BGLPSGSRITLHYG 75  
 DB 1 EFVLQNHADIVSCVAMVFLGLLMEITAKVSIIFVTLQYNVTLPAEQATESAFLYYG 60  
 QY 76 VKDLATVFFMLVAIIHATIOEYVLDKLSRRLQTKQKQNKLEAGOLSVFIIVSGIWG 135  
 DB 61 IKDLATVFFMLVAIIHATIOEYVLDKLSRRLQTKQKQNKLEAGOLSAFYLFSCIWG 120  
 QY 136 MTLASENCLSDPTLLWKSQPHNMFTQMKFFYSQLAYWPHSPPELYFQKVRKQDIPGQ 195  
 DB 121 TFLISENVISDPTILRAYPHNLMTQMKFFYSQLAYWPHAFPELYFQTKKEDIPRO 180  
 QY 196 LIYIGLHFLHGGAYLLNLHGLLLMLHYAVELLSSVCSLLYFGDERYOKGLSLMPIV 255  
 DB 181 LYVIIGLYLPHAGAYLLNLHGLVLLYHYFVEFLHISRLFYFSDKEYOKGFSWLAVL 240  
 QY 256 FTSGRLVTLIVSVTVVGLHLA-GTNRGNALSGNNVLAIAVLSSSCSTQVYITWTLT 314  
 DB 241 FVLGRLTLILSVLVGFLGARAENQKLDFTGNFNLAVRIAVLASICITQAFMMKKFI 300  
 QY 315 TWVLQRLWEDANLHVCGKRK-----RSRS-RKGTENGVE---NPNRIDSPPKKKEA 362  
 DB 301 NFOLRRWREHSAFOAPAVKKPPVTKGSRXXKGTENGVTGVTSGNAGDSRNRKES 357

RESULT 4  
 Y557\_HUMAN  
 ID Y557\_HUMAN STANDARD; PRT; 370 AA.  
 AC Q15035;  
 DT 16-OCT-2001 (Rel. 40, Created)  
 DT 16-OCT-2001 (Rel. 40, Last sequence update)  
 DE Hypothetical protein KIAA0057.  
 GN KIAA0057.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 OX NCBI\_TaxID=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Bone marrow;  
 RX MEDLINE=96051398; PubMed=7584044;  
 RA Nomura N., Nagase T., Miyajima N., Sazuka T., Tanaka A., Sato S.,  
 RA Seki N., Kawarabayashi Y., Ishikawa K.-I., Tabata S.;  
 RT "Prediction of the coding sequences of unidentified human genes. II.  
 RT The coding sequences of 40 new genes (KIAA0041-KIAA0080) deduced by  
 RT analysis of cDNA clones from human cell line KG-1.";  
 RL DNA Res. 1:223-229(1994).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RA Tracey A.;  
 RL Submitted (DEC-1999) to the EMBL/GenBank/DBJ databases.  
 CC -!- SUBCELLULAR LOCATION: Integral membrane protein (Potential).  
 CC -!- SIMILARITY: BELONGS TO THE LASS1 FAMILY.  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC EMBL: D31762; BAA06540.1; -  
 DR EMBL: AL049611; CAB71119.1; -  
 KW Hypothetical protein; Transmembrane.  
 FT TRANSMEM 23 43 POTENTIAL.  
 FT TRANSMEM 76 96 POTENTIAL.  
 FT TRANSMEM 120 140 POTENTIAL.  
 FT TRANSMEM 160 180 POTENTIAL.  
 FT TRANSMEM 199 219 POTENTIAL.  
 FT TRANSMEM 251 271 POTENTIAL.  
 FT TRANSMEM 288 308 POTENTIAL.  
 SQ SEQUENCE 370 AA; 43327 MW; 9B5183F1A3D45366 CRC64;  
 Query Match 39.9%; Score 754; DB 1; Length 370;  
 Best Local Similarity 42.4%; Pred. No. 1.1e-55;  
 Matches 158; Conservative 63; Mismatches 132; Indels 20; Gaps 5;  
 QY 5 KKNARNPVLVSHEFVNQADVSCVGMFFVLGLMFEPTAEMSIIVFTLQHGVVVPAEGL 64  
 DB 4 RRTTSYPLFSQEFVIHNDHAGIDGCLVCLVGLMFEVTAFTAKTAFILPQYNISVPT 60  
 QY 65 PGSGRTL-YHYGVKDLATVFFMLVAIIHATIOEYVLDKLSRRLQTKQKQNKLEAGQ 123  
 DB 61 -ADSETVHYHGPDLVTLFYIFITILHNAVQOEYILDKISKRHLHSKVHSPNSEQG 119  
 QY 124 LSVFIVSGIWMITLASENCLSDPTLLWKSQPHNMFTQMKFFYSQLAYWPHSPPELY 183  
 DB 120 LVVFHTSVINCFFVTVTEGYLTNPSLWEDYPHVLPFQVQKFFYLQCLAYWLHALPELY 179  
 QY 184 FOKVRKQDIPGOLYIGLHFLHGGAYLLNLHGLLLMLHYAVELLSSVCSLLYFGDE 243  
 DB 180 FOKVRKEETPQOLYICLYLPHAGAYLLNLHGLVLLYHYFVEFLHISRLFYFSTEFHTARLFYADE 239  
 QY 244 RYQKGLSLWPTVIFISGRVTLIVSVTVVGLHLA-GTNRGNALSGNNVLAIAVLSSS 302

Db	240	NNEKLFSAWAAVGYTRFLITLAVLAIGFLARMENQAFDEKGNFWTLFCRLCVLLLV	290	
Qy	303	CSQIVYITWTLYWLRQWLEDANLHVCGRKR	-----SRSRKGTENGVENPN	350
Db	300	CAQAQWLMRRFTHSQLRHWRWNEQ--SAKRRVPATPRLPARLIKRESGYHENGVVKA	E	357
Qy	351	RIDSPKKKEKAP	363	
Db	358	NGTSPRTKRLKSP	370	
RESULT 5				
LAGL YEAST				
ID	LAGL YEAST	STANDARD;	PRT;	411 AA.
AC	C38703;			
DT	01-FEB-1995 (Rel. 31, Created)			
DT	01-FEB-1995 (Rel. 31, Last sequence update)			
DE	16-OCT-2001 (Rel. 40, Last annotation update)			
DE	Longevity-assurance protein 1 (Longevity assurance factor 1).			
GN	LAGL OR YHLO03C.			
OS	Saccharomyces cerevisiae (Baker's yeast).			
OC	Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;			
OC	Saccharomycetales; Saccharomycetaceae; Saccharomycetes.			
OC	NCBI_TaxID=4932;			
RN	[1]			
RN	SEQUENCE FROM N.A.			
RP	STRAIN=X2180-1A;			
RC	MEDLINE=94253121; PubMed=8195187;			
RC	D'Mello N.P., Childress A.M., Franklin D.S., Kale S.P.,			
RA	Pinawadi C., Jazwinski S.M.;			
RA	"Cloning and characterization of LAGL, a longevity-assurance gene in			
RT	yeast.;			
RT	J. Biol. Chem. 269:15451-15459(1994).			
RN	[2]			
RN	SEQUENCE FROM N.A.			
RP	STRAIN=S288C / AB972;			
RC	MEDLINE=94378003; PubMed=8091229;			
RC	Johnston M., Andrews S., Brinkman R., Cooper J., Ding H., Dover J.,			
RA	Du Z., Favello A., Fulton L., Gattung S., Geisel C., Kirsten J.,			
RA	Kucaba T., Hillier L., Jier M., Johnston L., Langston Y.,			
RA	Latreille P., Louis E.J., Macri C., Mardis E., Menezes S., Mouser L.,			
RA	Nhan M., Rifkin L., Riles L., St Peter H., Trevaskis E., Vaughan K.,			
RA	Vignati D., Wilcox L., Wohlman P., Waterston R., Wilson R.,			
RA	Vaudin M.;			
RT	"Complete nucleotide sequence of Saccharomyces cerevisiae chromosome			
RT	VIII.;"			
RL	Science 265:2077-2082(1994).			
CC	-!- FUNCTION: INVOLVED IN THE AGING PROCESS. DELETION OF LAG1 RESULTS			
CC	IN A PROMINENT INCREASE (APPROXIMATELY 50%) IN MEAN AND IN			
CC	MAXIMUM LIFE SPAN.			
CC	-!- SUBCELLULAR LOCATION: Integral membrane protein (Potential).			
CC	-!- SIMILARITY: BELONGS TO THE LASSI FAMILY.			
CC	-----			
CC	This SWISS-PROT entry is copyright. It is produced through a collabora-			
CC	tion between the Swiss Institute of Bioinformatics and the EMBL outstat-			
CC	the European Bioinformatics Institute. There are no restrictions on			
CC	use by non-profit institutions as long as its content is in no			
CC	modified and this statement is not removed. Usage by and for commo-			
CC	entities requires a license agreement (See <a href="http://www.isb-sib.ch/annou">http://www.isb-sib.ch/annou</a>			
CC	or send an email to <a href="mailto:license@isb-sib.ch">license@isb-sib.ch</a> ).			
CC	-----			
DR	EMBL; U08133; AAA21579.1;			
DR	EMBL; U10555; AAB68429.1;			
DR	PIR; S46800; S46800.			
DR	SGD; S000095; LAG1.			
KW	Transmembrane.			
FT	TRANSMEM 82			POTENTIAL.
FT	TRANSMEM 135			POTENTIAL.
FT	TRANSMEM 177			POTENTIAL.
FT	TRANSMEM 212			POTENTIAL.
FT	TRANSMEM 252			POTENTIAL.
FT	TRANSMEM 297			POTENTIAL.
FT	TRANSMEM 317			POTENTIAL.
FT	TRANSMEM 356			POTENTIAL.



CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----

DR EMBL; X61398; CAA43670.1; -  
 DR EMBL; S59773; AAC60549.1; -  
 DR EMBL; Z28008; CAA81843.1; -  
 DR PIR; S30134; S30134.  
 DR SGD; S0001491; YKL008C.  
 KW Hypothetical protein; Transmembrane.  
 FT TRANSMEM 82 102 POTENTIAL.  
 FT TRANSMEM 135 155 POTENTIAL.  
 FT TRANSMEM 173 193 POTENTIAL.  
 FT TRANSMEM 260 280 POTENTIAL.  
 FT TRANSMEM 297 317 POTENTIAL.  
 FT TRANSMEM 356 376 POTENTIAL.  
 SQ SEQUENCE 418 AA; 48992 MW; 7691BA623AC0460A CRC64;

Query Match 7.2%; Score 137; DB 1; Length 418;  
 Best Local Similarity 23.5%; Pred. No. 0.00034;  
 Matches 61; Conservative 44; Mismatches 93; Indels 62; Gaps 11;

QY 67 GSTRTHYGVKDLATVFFVYMLVAIIHATIQEYVLDKLSRRQLTKGQKNLEAGOLSV 126  
 DB 121 GDTNAYGKINDLCFVFFYVFFTEFLREFLDVIRPAIRLHVTSHKRIKRIEOMYAI 180  
 QY 127 FYI-VSGIWMIL-ASENCLESDPTLLMKSPHNMFTQMKFFYSOLAYW----- 175  
 DB 181 FYTVSGPGIYCMYHSDLWFFNTAMRTYDFTNPLFKFYLGQAQFAWQAACILVL 240  
 QY 176 -----FHSFPEL-----YFQVKQDIPQGLIYIGLPHIGGAYLLXNH 216  
 DB 241 QLEKPKRDNELTFHHVTLVLLWSYVFFHTKMGLP---IVITMDVDFLLSFKTLNY 297  
 QY 217 L-GLL-----LLMHYA-VELLSSVCS-----LIYFGDERYQKGLSLWPI 254  
 DB 298 LSGLAFFSFAIVVAMHYLRHYINLKILMSVLTOFRTEGNYVLNFAQQYKWCWISL-PI 356  
 QY 255 VFISGRVTLIVSVTVGLH 274  
 DB 357 VFV-----LIGALQVLNLY 370

RESULT 7  
 ID YHFXF SCHPO STANDARD; PRT; 384 AA.  
 AC O59735;  
 DT 16-OCT-2001 (Rel. 40, Created)  
 DT 16-OCT-2001 (Rel. 40, Last sequence update)  
 DE Hypothetical 42.4 kDa protein C3E17.15c in chromosome II.  
 GN SPBG3E7.15C OR SPBC4F6.02C.  
 OS Schizosaccharomyces pombe (Fission yeast).  
 OC Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;  
 OC Schizosaccharomycetales; Schizosaccharomycetaceae;  
 OC NCBI\_TaxID=4896;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=972;  
 RA Lyne M., Wood V., Rajandream M.A., Barrell B.G., Brown D.,  
 RA Churcher C.M.;  
 RL Submitted (MAY-1998) to the EMBL/GenBank/DBJ databases.  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=972;  
 RA William R., Rajandream M.A., Barrell B.G., Skelton J., Churcher C.M.;  
 RL Submitted (SEP-1998) to the EMBL/GenBank/DBJ databases.  
 CC -!- SUBCELLULAR LOCATION: Integral membrane protein (Potential).

CC -!- SIMILARITY: BELONGS TO THE LASSI FAMILY.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----

DR EMBL; AL023534; CAA19018.2; -  
 DR EMBL; AL031534; CAA20722.2; -  
 KW Hypothetical protein; Transmembrane.  
 FT TRANSMEM 59 79 POTENTIAL.  
 FT TRANSMEM 104 124 POTENTIAL.  
 FT TRANSMEM 155 175 POTENTIAL.  
 FT TRANSMEM 195 215 POTENTIAL.  
 FT TRANSMEM 233 253 POTENTIAL.  
 FT TRANSMEM 270 290 POTENTIAL.  
 FT TRANSMEM 329 349 POTENTIAL.  
 SQ SEQUENCE 384 AA; 45335 MW; 82FCF8EA6638849A CRC64;

Query Match 7.0%; Score 133; DB 1; Length 384;  
 Best Local Similarity 20.5%; Pred. No. 0.00067;  
 Matches 48; Conservative 45; Mismatches 117; Indels 24; Gaps 7;

QY 41 EGTAEISVFLTL-----OHGVVPAEGLPSGSRTHYGVKDLATVFFVYML 87  
 DB 56 EKTWIVPLILLTLGVWFPVNPNGYKIGFL-SYPIGTNPAQYKGKRLDIAFLFYAL 114  
 QY 88 VAIITHATQEVLDKLSRRQL-TKGKONLEAGOLSVFYIVSGIWMIL-ASENCLE 145  
 DB 115 FTFCTREFIMQIIAIGHFNRAPAKLRFEEOAYCLFYTVMSGLVYVYKTPMWF 174  
 QY 146 SDPTLLWKSPHNMFTQMKFFYSOLAYWFSFPELYFQ--KVRKQDIPQGLIYIGLH 203  
 DB 175 FNTDAFWEYEPHYHVGSPKAFYIEAAWVIOQALVLIQLEKPRK-DFKELVVHHITL 233  
 QY 204 FHIGGAYLLYLNHLLMLHYAVELLSVCSLLYFGDERYQKGLSLWPIVFI 257  
 DB 234 LLTGLSYFFHTWIGLAVFITMDTSDIWLKSLCL-----NYVNTVIVYPIFVI 282

RESULT 8  
 ID LAGL SCHPO STANDARD; PRT; 390 AA.  
 AC P78970; O13860;  
 DT 01-NOV-1997 (Rel. 35, Created)  
 DT 15-DEC-1998 (Rel. 37, Last sequence update)  
 DT 16-OCT-2001 (Rel. 40, Last annotation update)  
 DE Longevity-assurance protein 1 (Longevity assurance factor 1).  
 GN LAG1 OR SPAC1A6.09C.  
 OS Schizosaccharomyces pombe (Fission yeast).  
 OC Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;  
 OC Schizosaccharomycetales; Schizosaccharomycetaceae;  
 OC NCBI\_TaxID=4896;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Chanda E.R., Lingner C., Ko Z., Young P.G.;  
 RL Submitted (NOV-1996) to the EMBL/GenBank/DBJ databases.  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=972;  
 RA Churcher C.M., Barrell B.G., Rajandream M.A., Walsh S.V., Wood V.;  
 RL Submitted (SEP-1997) to the EMBL/GenBank/DBJ databases.  
 CC -!- FUNCTION: INVOLVED IN THE AGING PROCESS (BY SIMILARITY).  
 CC -!- SUBCELLULAR LOCATION: Integral membrane protein (Potential).  
 CC -!- SIMILARITY: BELONGS TO THE LASSI FAMILY.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -

CC The European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

DR EMBL; U76608; AAB19113.1; -;  
DR EMBL; 299258; CAB16359.1; -;  
KW Transmembrane.  
FT DOMAIN 13 16 POLY-SER.  
FT TRANSMEM 63 83 POTENTIAL.  
FT TRANSMEM 117 137 POTENTIAL.  
FT TRANSMEM 155 175 POTENTIAL.  
FT TRANSMEM 199 219 POTENTIAL.  
FT TRANSMEM 231 251 POTENTIAL.  
FT TRANSMEM 276 296 POTENTIAL.  
FT TRANSMEM 336 356 POTENTIAL.  
FT CONFLICT 379 390 EDEASSTNEDK -> GRRGERNE (IN REF. 1).  
SQ SEQUENCE 390 AA; 45668 MW; DC00FB5C2D2F22CC CRC64;

Query Match 6.68; Score 125; DB 1; Length 390;  
Best Local Similarity 23.98; Pred. No. 0.0032;  
Matches 57; Conservative 35; Mismatches 100; Indels 46; Gaps 11;

QY 65 PSGSRTLYHYGVKDLATVFFYMLVAIIHATIQEYVLDKLSRRLQLTKGKON-----KLN 119  
DB 104 PDGS---YKGPDKACFFIFWVFTAFRIVMDYVF-----RPVFWGVRNRKVIIFRC 156  
QY 120 EAGQLSVFYVSGIWMIIASENCSDPTLLKWSQPHNMFTQMKPFYISQLAYWPHSF 179  
DB 157 EOGYSFFYLFCFWFLGLYIRSSYWSNEEKLFDYPOYVMSPLFKAYYLQGLFWLQOI 216  
QY 180 PELYFQVRKOD-----IPGQLIVI--GLHLFHGGAYLALNHLGLLL-----ML 224  
DB 217 LVHLHQ-RRADWQMFHHVTCALILSYGFNFRVGN-ILYIFDLSYILSGKML 274  
QY 225 HYAVELLSSVCLLYFGDERYQKGLSLWPIYVIFSGRLVTLIVSVTV-----GLHL 275  
DB 275 KYLG--FGKICDYLF-----GIFVASWYSRHYLFSLKILRVVTVNAPEIIGPHL 322

## RESULT 9

ID GLPT\_HUMAN STANDARD; PRT; 533 AA.  
AC P57057; Q9HAQ1;  
DT 16-OCT-2001 (Rel. 40, Created)  
DT 16-OCT-2001 (Rel. 40, Last sequence update)  
DT 16-OCT-2001 (Rel. 40, Last annotation update)  
DE Glycerol-3-phosphate transporter (G-3-P transporter) (G-3-P  
DE permease).  
GN SLC37A1.  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
OX NCBI\_TaxID=9606;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC TISSUE=Fetal brain;  
RA Bartoloni L., Wattenhofer M., Kudoh J., Kawasaki K., Rossier C.,  
RA Shimizu N., Scott H.S., Antonarakis S.E.;  
RT "Identification and characterization of the human glycerol 3-phosphate  
RT permease gene (SLC37A1) mapping to 21q22.3; expression pattern,  
RT genomic structure and cDNA sequence."  
RL Submitted (MAY-2000) to the EMBL/GenBank/DBJ databases.  
RN [2]  
RP SEQUENCE FROM N.A.  
RA Solans A., Estivill X., de la Luna S.;  
RT "Cloning and characterization of human glycerol 3-phosphate permease  
RT gene (SLC37A1)."  
RL Submitted (OCT-2000) to the EMBL/GenBank/DBJ databases.  
CC -1- SUBCELLULAR LOCATION: Integral membrane protein (Potential).  
CC -1- SIMILARITY: BELONGS TO THE SLC37A FAMILY OF TRANSPORTERS.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

DR EMBL; AJ269529; CAB87248.1; -;  
DR EMBL; AJ277912; CAB91985.1; -;  
DR EMBL; AJ277913; CAB91986.1; -;  
DR EMBL; AF311320; AAG29853.1; -;  
DR InterPro; IPR000849; GLPT\_transporter.  
DR PROSITE; PS00942; GLPT; FALSE\_NEG.  
KW Transmembrane; Transport; Sugar transport.  
FT TRANSMEM 18 38 POTENTIAL.  
FT TRANSMEM 100 120 POTENTIAL.  
FT TRANSMEM 129 149 POTENTIAL.  
FT TRANSMEM 157 177 POTENTIAL.  
FT TRANSMEM 222 242 POTENTIAL.  
FT TRANSMEM 304 324 POTENTIAL.  
FT TRANSMEM 334 354 POTENTIAL.  
FT TRANSMEM 366 386 POTENTIAL.  
FT TRANSMEM 394 414 POTENTIAL.  
FT TRANSMEM 423 443 POTENTIAL.  
FT TRANSMEM 466 486 POTENTIAL.  
FT TRANSMEM 490 510 POTENTIAL.  
FT CONFLICT 223 223 F -> L (IN REF. 2).  
FT CONFLICT 303 303 H -> Q (IN REF. 2).  
SQ SEQUENCE 533 AA; 57662 MW; EA888FE2942380A2 CRC64;

Query Match 5.28; Score 97.5; DB 1; Length 533;  
Best Local Similarity 20.18; Pred. No. 0.9; Mismatches 108; Indels 107; Gaps 17;  
Matches 69; Conservative 59;

QY 26 MVSCVGMFF-----VLGLMEFG---TAEMSIIVFLTQHGIVVPAE 62  
DB 176 VVTCLGNWFGKRRGLMGVWNSHTSVGNILGSLIAGYVWSTCWGLSFVV--PGAIVAAM 233  
QY 63 GLPSSRTLYHYGVKDLATVFFYMLV-----AIIHATIQEYVLDKLSRRLQLTK 112  
DB 234 GI-----VCLFLIEHPNDVRCSTLTVTHSKGYENGNTNRLRLQKILK 276  
QY 113 GKONK-----LNEA-GOLSVFYIV-----SGIWMIIILASENCLSDPTLLKWSQPH 157  
DB 277 SEKNPPLDEPMOCLLLSDGKSIHPNHVILPGDGGSGTAASFTGALKIP-----327  
QY 158 NMWTFQMKFFYISQLAYWPHSFPELYFQKVRKQDI--PGQLIYIGLHFLHIGGAYLLYN 215  
DB 328 GVIEFSCLLLFAKLVSYTFLFWLPYITNVDHLDAKAGEL-----STLFDVGGIF-----378  
QY 216 HGLGLLLMLHYAVELLSSVCSL-----LYFGDERYQKGL-SLWPIVFIISGRV----262  
DB 379 -GGILAGVISDRLEKRASTCGMLLLLAAPTLYIFSTIKMGLEATIAMLLSGALVSGPY 437  
QY 263 TLIVSVTVVGLHLAGTNRNGNALSGNVNVLAKIAVLSSSCSI 305  
DB 438 TLITTAVSADL---GTHK---SLKGNALHSIVTALIDTGTGV 474

## RESULT 10

ID CFTR\_RABIT STANDARD; PRT; 1450 AA.  
AC Q00554;  
DT 01-FEB-1994 (Rel. 28, Created)  
DT 01-OCT-1996 (Rel. 34, Last sequence update)  
DT 01-MAR-2002 (Rel. 41, Last annotation update)  
DE Cystic fibrosis transmembrane conductance regulator (CFTR) (CAMP-  
DE dependent chloride channel).  
GN ABC7 OR CFTR.  
OS Oryctolagus cuniculus (Rabbit).

[illegible]

CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

DR EMBL; Y13855; CA74167.1; -;  
DR EMBL; AJ235271; CA14862.1; -;  
DR HSSP; P98002; IARI.  
DR InterPro; IPR000883; COX1.  
DR Pfam; PF00115; COX1; 1.  
DR PRINTS; PR01165; CYCOXIDASE1.  
DR PROSITE; PS00077; COX1; 1.  
KW Oxidoreductase; Heme; Copper; Transmembrane; Respiratory chain;  
KW Complete proteome.  
FT TRANSMEM 35 55 POTENTIAL.  
FT TRANSMEM 76 96 POTENTIAL.  
FT TRANSMEM 97 117 POTENTIAL.  
FT TRANSMEM 120 140 POTENTIAL.  
FT TRANSMEM 165 185 POTENTIAL.  
FT TRANSMEM 202 222 POTENTIAL.  
FT TRANSMEM 254 274 POTENTIAL.  
FT TRANSMEM 286 306 POTENTIAL.  
FT TRANSMEM 320 340 POTENTIAL.  
FT TRANSMEM 357 377 POTENTIAL.  
FT TRANSMEM 396 416 POTENTIAL.  
FT TRANSMEM 433 453 POTENTIAL.  
FT TRANSMEM 475 495 POTENTIAL.  
FT METAL 81 81 IRON (HEME A) (PROBABLE).  
FT METAL 260 260 COPPER B (PROBABLE).  
FT METAL 264 264 COPPER B (PROBABLE).  
FT METAL 309 309 COPPER B (PROBABLE).  
FT METAL 310 310 COPPER B (PROBABLE).  
FT METAL 395 395 IRON (HEME A3) (PROBABLE).  
FT METAL 397 397 IRON (HEME A) (PROBABLE).  
SQ SEQUENCE 534 AA; 59261 MW; 326685753548C3E CRC64;

Query Match 5.1%; Score 96.5; DB 1; Length 534;  
Best Local Similarity 18.6%; Pred. No. 1.1;  
Matches 65; Conservative 57; Mismatches 107; Indels 121; Gaps 19;  
QY 14 LSHPEWQNHADMSCVGM--FFVLGLMFEGTAEMSVFL-----TLQHGCVV 59  
DB 67 LNHFNLYNLVTHAIIMVFMMIPALFSGFGNYFVPLLTGAPDMAPRLNNISEWLLI 126  
QY 60 PA-----EGLPGSSRTLY-----HYGVKDLATVEFFYMLVAIIHATIQEYVL 101  
DB 127 PAFLLISSITDGGGTGWTLYPPLSNLNGHTGAADVVAIF-----SLHITGLSSIL 179  
QY 102 DKLRRRLQ-----TKGKONKLNAGQSVFVIVSGIWMGMIILASENCLSDPTLLWKSQPH 157  
DB 180 GSINLIVTFNNRTPG-----NGLKMPLEFVNSILVTAFLIILAMPVL---SGAI 226  
QY 158 NMW---TFQMKFF-----YISQLAYWFHSPPELYF-----QKV---RK-- 199  
DB 227 TMLITDRNFGFTTFKPDGGDPLLFQHLFFWFGHPHYIIVILPGFVGIVSQVISTFRRKPI 286  
QY 190 ---QDIPQQLIYIGL-----HLFHIGGAY---LLYLNHLGLLLMLHYAVELLSSVCS 236  
DB 287 FGXOGMGAMVILGFGVFIWAHMFVGLSYNALIVFT-AGTWIAVPTGKIFSWIA- 344  
QY 237 LLYFGDERYQKGLSLW-----PIVFISGRVLTVLSVVTVGLHLAQT 278  
DB 345 -----TWGSGITFPTPMLFAIGFIILFTIGGVT-GIILSNS 380  
RESULT 12  
YDIK\_ECOLI STANDARD; PRT; 370 AA.  
AC P77175;  
DT 16-OCT-2001 (Rel. 40, Created)  
DT 16-OCT-2001 (Rel. 40, Last sequence update)

DT 16-OCT-2001 (Rel. 40, Last annotation update)  
GN Hypothetical protein ydiK.  
DE YDIK OR B1688.  
OS Escherichia coli.  
OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;  
OC Escherichia  
ON NCBI\_TaxID=562;  
RX SEQUENCE FROM N.A.  
RP STRAIN-K12 / MG1655;  
RC MEDLINE-97426617; PubMed-9278503;  
RA Blattner F.R., Plunkett G. III, Bloch C.A., Perna N.T., Burland V.,  
RA Riley M., Collado-Vides J., Glasner F.D., Rode C.K., Mayhew G.F.,  
RA Gregor J., Davis N.W., Kirkpatrick H.A., Goeden M.A., Rose D.J.,  
RA Mau B., Shao Y.;  
RT "The complete genome sequence of Escherichia coli K-12.";  
RL Science 277:1453-1474(1997).  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN-K12;  
RX MEDLINE-97251357; PubMed-9097039;  
RA Aiba H., Baba T., Fujita K., Hayashi K., Inada T., Isono K., Itoh T.,  
RA Kasai H., Kashimoto K., Kimura S., Kitakawa M., Kitagawa M.,  
RA Makino K., Miki T., Mizobuchi K., Mori H., Mori T., Motomura K.,  
RA Nakade S., Nakamura Y., Nashimoto H., Nishio Y., Oshima T., Saito N.,  
RA Sampei G., Seki Y., Sivasubram S., Tagami H., Takeda J.,  
RA Takemoto K., Takeuchi Y., Wada C., Yamamoto Y., Horiuchi T.;  
RT "A 570-kb DNA sequence of the Escherichia coli K-12 genome  
corresponding to the 28.0-40.1 min region on the linkage map.";  
RL DNA Res. 3:363-377(1996).  
CC - SUBCELLULAR LOCATION: Integral membrane protein (Potential).  
CC - SIMILARITY: BELONGS TO THE UPF0118 (PERM) FAMILY.  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

EMBL; AE000264; AAC74758.1; -;  
EMBL; D90811; BAA15450.1; -;  
DR EcoGene; EG13970; ydiK.  
DR InterPro; IPR002549; UPF0118.  
DR Pfam; PF01594; UPF0118; 1.  
KW Hypothetical protein; Transmembrane; Complete proteome.  
FT TRANSMEM 18 38 POTENTIAL.  
FT TRANSMEM 40 60 POTENTIAL.  
FT TRANSMEM 66 86 POTENTIAL.  
FT TRANSMEM 107 127 POTENTIAL.  
FT TRANSMEM 157 177 POTENTIAL.  
FT TRANSMEM 214 234 POTENTIAL.  
FT TRANSMEM 249 269 POTENTIAL.  
FT TRANSMEM 275 295 POTENTIAL.  
FT TRANSMEM 309 329 POTENTIAL.  
SQ SEQUENCE 370 AA; 39841 MW; 14B7AD79D8C56248 CRC64;  
Query Match 5.0%; Score 95; DB 1; Length 370;  
Best Local Similarity 19.5%; Pred. No. 0.94;  
Matches 67; Conservative 50; Mismatches 86; Indels 140; Gaps 16;  
QY 26 MVSCVGMF--FVLGLMFEGTAEMSI--VFLTIQHGVVVPVPAEGLPSGSRTLHYGVKDLAT 81  
DB 24 IVACLWIQVPFIFGFAWAGTVVIAWVPLLRLQK-----IMFGRSLAVLVMTLLV 75  
QY 82 VFFYMLVAIIHATIQEYVLDKLSRRLQLTGKONKLNAGQSVFVIVSGIWMGMIILAS 141  
DB 76 MVFIPIALLVNS-----IVDGGSPLIKAI 101  
QY 142 ENCLSDPTLLWKSQPHNMFTFMKFFIISQLAYWFHSPPELYFQKVRKQDIPGQLIYIGL 201

Db 102 SGDMTLPDLA-----WLNTIP-----VIGAKLIAGW 127

Qy 202 H-LFHIGGAYLL-----YLN-----HLGILLMLHYAVELLSSVCSLLYFGDER 244

Db 128 HNLDMGTAIMAKVRPIYIGITTTTFWVGQAHHG--RFMVHCALMLFS--ALLYRGEQ 183

Qy 245 YQKG-----LSLWPIVPIIS--GRILVTLIVSVVTVGLHAGTNRNGNAL 285

Db 184 VAQGIHFATRLAGVRGDAAVLLAAQRAVALGVVVVTVLQVAVLGGI-----GLAV 235

Qy 286 SGNNVLAIAVLLSSCSIQ-----VYITWTLTIV 317

Db 236 SG-VPYATLTLVLMILSLVQLGCLPLVLPAILWLTWTGDTTW 277

RESULT 13

SOAL\_MOUSE

ID SOAL\_MOUSE STANDARD; PRT; 540 AA.

AC Q61263; Q64180;

DT 16-OCT-2001 (Rel. 40, Created)

DT 16-OCT-2001 (Rel. 40, Last sequence update)

DE Sterol O-acyltransferase 1 (EC 2.3.1.26) (Cholesterol acyltransferase 1) (Acyl coenzyme A:cholesterol acyltransferase 1) (ACAT-1).

GN SOAT1 OR ACAT.

OS Mus musculus (Mouse).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

OX NCBI\_TaxID=10090;

RN [1]

RP SEQUENCE FROM N.A.

RC TISSUE=Liver;

RX MEDLINE=96064687; PubMed=7592824;

RA Uselman P.J., Oka K., Sullivan M.C., Chang T.-Y., Chang C.C.Y., Chan L.;

RT "Tissue-specific expression and cholesterol regulation of acylcoenzyme A:cholesterol acyltransferase (ACAT) in mice. Molecular cloning of mouse ACAT cDNA, chromosomal localization, and regulation of ACAT in vivo and in vitro.";

RT J. Biol. Chem. 270:26192-26201(1995).

RN [2]

RP SEQUENCE FROM N.A.

RX MEDLINE=96158986; PubMed=8579615;

RA Green S., Steinberg D., Quehenberger O.;

RT "Cloning and expression in Xenopus oocytes of a mouse homologue of the human acylcoenzyme A: cholesterol acyltransferase and its potential role in metabolism of oxidized LDL.";

RT Biochem. Biophys. Res. Commun. 218:924-929(1996).

RL -1- FUNCTION: CATALYZES THE FORMATION OF FATTY ACID-CHOLESTEROL ESTERS. PLAYS A ROLE IN LIPOPROTEIN ASSEMBLY AND DIETARY CHOLESTEROL ABSORPTION.

CC -1- CATALYTIC ACTIVITY: Acyl-CoA + cholesterol = CoA + cholesterol ester.

CC -1- SUBCELLULAR LOCATION: Integral membrane protein. Endoplasmic reticulum.

CC -1- SIMILARITY: BELONGS TO THE STEROL O-ACYLTRANSFERASE FAMILY.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

CC EMBL: L42293; AAC42075.1; -

DR EMBL: S81092; AAB36050.1; -

DR MGD: MGI:104665; Soatl.

DR InterPro: IPR002688; ACAT.

DR Pfam: PF01800; ACAT; 1.

KW Transferase; Acyltransferase; Transmembrane; Endoplasmic reticulum; Cholesterol metabolism.

FT TRANSMEM 132 152

POTENTIAL.

FT TRANSMEM 311 331

FT TRANSMEM 352 372

FT TRANSMEM 460 480

FT TRANSMEM 488 508

FT CONFLICT 195 195

SQ SEQUENCE 540 AA; 63739 MW; 8EF900C8BCDF73C0 CRC64;

Query Match

Best Local Similarity 5.0%; Score 94.5; DB 1; Length 540;

Matches 45; Conservative 34; Mismatches 93; Indels 37; Gaps 6;

Qy 48 IVFLTLQHGCV-----VVPAGELPSGSRITLYHGVKDLATVFFYMLVAIIHATIEYV 100

Db 215 LLFLVFLQVLGVFTVYVYVLAITLPPASR-----FILILEQIRLIMKAHSFV 261

Qy 101 LDKLSRRRLQLTKGKQNK--LNEAGQLSVFYIVSGIWMGMIILASENCLSDPTLLKKSOPHN 158

Db 262 RENIPRVLNAAKEKSSKDPLPTVNOQLYF-----LEAPTLLIYRDNYPRTPVWR--GYV 313

Qy 159 MWTQMKEFFYISQAYWPHSPPELYFQKVRKODIPGQLIYIGLHLFHIGGAYLLNLHLG 218

Db 314 AMQFLQVGCFLFYVYIIFERLCAPLFRNIKQEPFSARVLYCVFNSILPGVLILFLSFFA 373

Qy 219 LLLMLHYAVELLSSVCSLLYFGDERYQK 247

Db 374 FLHCWLNFAEMLR-----FGDRMFYK 395

RESULT 14

QOXM\_SULAC

ID QOXM\_SULAC STANDARD; PRT; 788 AA.

AC P39481;

DT 01-FEB-1995 (Rel. 31, Created)

DT 01-FEB-1995 (Rel. 31, Last sequence update)

DT 30-MAY-2000 (Rel. 39, Last annotation update)

DE Quinol oxidase polypeptide I/III (EC 1.9.3.-).

GN SOXM.

OS Sulfolobus acidocaldarius.

OC Archaea; Crenarchaeota; Sulfolobales; Sulfolobaceae; Sulfolobus.

OX NCBI\_TaxID=2285;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=ATCC 33909 / NCIB 11770 / DSM 639;

RX MEDLINE=94357214; PubMed=8076636;

RA Luebben M., Arnaud S., Castresana J., Warne A., Albracht S.P.J., Saraste M.;

RT "A second terminal oxidase in Sulfolobus acidocaldarius.";

RL Eur. J. Biochem. 224:151-159(1994).

CC -1- FUNCTION: TERMINAL OXIDASE IS THE COMPONENT OF THE RESPIRATORY CHAIN THAT CATALYZES THE REDUCTION OF OXYGEN TO WATER. SOXM FORMS THE FUNCTIONAL CORE OF THE ENZYME COMPLEX.

CC -1- FUNCTION: CATALYTIC SUBUNIT OF THE ENZYME. ELECTRONS ORIGINATING IN A QUINOL ARE TRANSFERRED TO THE BIMETALLIC CENTER OF SOXM FORMED BY A HEME AND COPPER B.

CC -1- CATALYTIC ACTIVITY: 4 ferrocytochrome c + O(2) = 4 ferrocytochrome c + 2 H(2)O.

CC -1- COFACTOR: TWO HEME GROUPS AND COPPER B.

CC -1- PATHWAY: TERMINAL STEP IN THE RESPIRATORY CHAIN.

CC -1- SUBUNIT: FORMS A COMPLEX WITH AT LEAST SOXC AND A 30 kDa RIESKE FE-S PROTEIN, BUT NEITHER WITH SOXA NOR SOXB.

CC -1- SUBCELLULAR LOCATION: Integral membrane protein.

CC -1- PTM: SOXM IS PROBABLY A PRECURSOR FORM OF SUBUNITS I AND III.

CC -1- SIMILARITY: IN THE N-TERMINAL SECTION; BELONGS TO THE HEME-COPPER RESPIRATORY OXIDASE FAMILY.

CC -1- SIMILARITY: IN THE C-TERMINAL SECTION; BELONGS TO THE CYTOCHROME C OXIDASE SUBUNIT 3 FAMILY.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial

entitles requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

```
CC EMBL; X73567; CAA51969.1;
CC HSSP; P00396; 10CC.
CC InterPro; IPR000883; COX1.
CC InterPro; IPR000298; CytC_oxdse_III.
CC Pfam; PF00115; COX1; 1.
CC Pfam; PF00510; COX3; 1.
CC PRINTS; P01165; CYCOXIDASE1.
CC PROSITE; PS00077; COX1; 1.
CC PROSITE; PS0253; COX3; 1.
CC Oxidoreductase; Heme; Copper; Transmembrane; Respiratory chain.
KW TRANSMEM 19 39
FT TRANSMEM 58 78
FT TRANSMEM 100 119
FT TRANSMEM 152 169
FT TRANSMEM 187 206
FT TRANSMEM 232 252
FT TRANSMEM 276 294
FT TRANSMEM 307 326
FT TRANSMEM 337 356
FT TRANSMEM 369 388
FT TRANSMEM 411 432
FT TRANSMEM 449 468
FT TRANSMEM 526 544
FT TRANSMEM 547 564
FT TRANSMEM 614 634
FT TRANSMEM 657 676
FT TRANSMEM 690 707
FT TRANSMEM 729 749
FT TRANSMEM 766 786
FT METAL 62 62
FT METAL 238 238
FT METAL 242 242
FT METAL 288 288
FT METAL 289 289
FT METAL 374 374
FT METAL 376 376
SQ SEQUENCE 788 AA; 87082 MW; 3939C16CDB8A08AD CRC64;

Query Match 5.0%; Score 94.5; DB 1; Length 788;
Best Local Similarity 19.5%; Pred. No. 2.5;
Matches 73; Conservative 57; Mismatches 98; Indels 147; Gaps 19;

Qy 28 SCVG-MFEVLGL--MFECTAEMSIIV-----FLTLQ-----HG-----VVVP-AE 62
Db 17 SDVGQYIVLGIVALLIGSVNAALIRQLSFNNAVYDYDAVTLHGIFMFFVWPLST 76
Qy 63 GLPS-----GSTRLYHYGVKDLATVFFYMLVAIIITHAQE-----YVLDKLSR 106
Db 77 GFANYLVRPMIGADHLYWPKINALS---FWMLVPAVILAATISPLLGAVDLGWYMYAPLSV 133
Qy 107 RLQLTGKQKONKNEAGOLSVPIYVSG-----IMGMII 138
Db 134 ETTVNYGLGTNL-----IQALISGLSSLTGTVNFWMTTKMKKVPYLMKPLFVWGFFT 188
Qy 139 LASENCLSDPTLLMKSQPHNNMTFMKFFIYSOL-----AYNWFHSF 179
Db 189 TAILMIAMPSTL-----TAGLVFAYLERLWGPFFDSALGSGSPVLWQQLFWFFGH 238
Qy 180 PELY-----FQVKRKQIDIPGOLYIGLHFLHFGGAYL-----LYLNHLGLLLL 222
Db 239 PEVYILIPAMGLVSELLPKMARREIFG---YTAIALSSIAIAFLSALGVMMHH--MFTA 293
Qy 223 MLHAVAVELLSSVCSL-----LYGEDRYOKGLSL---WPVIFTSGLVLT 263
Db 294 IDNTLVQIVSSATWTAIPSGVAVLNWATLYGGEIRYKTPPTILLISFIVMFLGIGITG 353
Qy 264 LIVSVVTVGLHLAGT 278
Db 354 VFFPLVPIDYALNGT 368
```

RESULT 15  
ATP6\_BUCAP STANDARD; PRT; 272 AA.  
AC O51878;  
DT 15-DEC-1998 (Rel. 37, Created)  
DT 15-DEC-1998 (Rel. 37, Last sequence update)  
DE 30-MAY-2000 (Rel. 39, Last annotation update)  
DT ATP synthase A chain (EC 3.6.3.14) (Protein 6).  
GN ATPB.  
OS Buchnera aphidicola (subsp. Schizaphis graminum).  
OC Bacteria; Proteobacteria; gamma subdivision; Buchnera.  
OX NCBI\_TaxID=98794;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=97361981; PubMed=9216881;  
RA Clark M.A., Baumann P.;  
RT "The (F1F0) ATP synthase of Buchnera aphidicola (endosymbiont of aphids): genetic analysis of the putative ATP operon.";  
RL Curr. Microbiol. 35:84-89(1997).  
RN [2]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=98184963; PubMed=9516544;  
RA Clark M.A., Baumann L., Baumann P.;  
RT "Sequence analysis of a 34.7-kb DNA segment from the genome of Buchnera aphidicola (endosymbiont of aphids) containing groEL, dnaA, the atp operon, gldA, and rho.";  
RL Curr. Microbiol. 36:158-163(1998).  
CC -!- FUNCTION: KEY COMPONENT OF THE PROTON CHANNEL; IT MAY PLAY A DIRECT ROLE IN THE TRANSLOCATION OF PROTONS ACROSS THE MEMBRANE (BY SIMILARITY).  
CC -!- SUBUNIT: F-TYPE ATPASES HAVE 2 COMPONENTS, CF(1) - THE CATALYTIC CORE - AND CF(0) - THE MEMBRANE PROTON CHANNEL. CF(1) HAS FIVE SUBUNITS: ALPHA(3), BETA(3), GAMMA(1), DELTA(1), EPSILON(1). CF(0) HAS THREE MAIN SUBUNITS: A, B AND C (BY SIMILARITY).  
CC -!- SUBCELLULAR LOCATION: Integral membrane protein (potential).  
CC -!- SIMILARITY: BELONGS TO THE ATPASE A CHAIN FAMILY.  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL Outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC EMBL; AF008210; AAC381116.1; -  
DR HSSP; P00855; 1C17.  
DR InterPro; IPR000568; ATP\_synt\_A.  
DR Pfam; PF00119; ATP\_synt\_A; 1.  
DR PROSITE; PS00449; ATPASE\_A; 1.  
KW Hydrogen ion transport; CF(0); Transmembrane.  
FT TRANSMEM 41 61 POTENTIAL.  
FT TRANSMEM 110 130 POTENTIAL.  
FT TRANSMEM 156 176 POTENTIAL.  
FT TRANSMEM 188 208 POTENTIAL.  
FT TRANSMEM 222 242 POTENTIAL.  
FT TRANSMEM 243 263 POTENTIAL.  
SQ SEQUENCE 272 AA; 31253 MW; DCD8C7D2C98C37CC CRC64;

Query Match 5.0%; Score 94; DB 1; Length 272;  
Best Local Similarity 23.5%; Pred. No. 0.8;  
Matches 65; Conservative 45; Mismatches 87; Indels 80; Gaps 16;

Qy 73 HYGVKDLATVFFYMLVAIIITHAQEIVLDKLSRRLQLTGKQKONKNEAGOLSVPIYVSG 132  
Db 38 HWVLNIDSIIISFLVGCFFLSIF--YTVAK-----KITGVNGLQASIELIFDIRSN 90  
Qy 133 IWGMIIASENCLSDP-----TLLWKSQPHNNMTFMKFF-YISQLAYWPHSPPELYFOKV 187  
Db 91 VKSM--YQGNPLIAPLSLTVFVFWFLNMLDLIPIDFFPFISE--RFFH-FPAM----- 140

```

Qy 188 RKQIPGOLIYIGLHIFHIGGAYLLYNHGLGLLLMLHYAVEL--LSSVCSLLYFGDERY 245
Db 141 --RIVPSADINITLSM-----SLGVFILILFYSVKMKGLIGFC----- 176
Qy 246 QKGLSL---WPIVFISG---RLVTLIVSVTVVCLHLAGTNRNCGNLSGN----- 288
Db 177 -KELTLOPFNHPVFIFNLELVLSLSPISGLRLEF-----GNMYSGEMIFILIAGLL 230
Qy 289 -----VNLAAKIAVLSSSCSIQVYITWTLTIVWL 318
Db 231 PWSQFFLNVPWAFHIL--IISIQAFIEMVLTIVL 265

```

Search completed: September 20, 2002, 14:39:18  
Job time: 245 sec







```
Query Match 40.3%; Score 762; DB 13; Length 371;
Best Local Similarity 41.2%; Pred. No. 2.5e-60;
Matches 159; Conservative 67; Mismatches 122; Indels 38; Gaps 7;

Qy 1 MGLRKKARPPVLSHEFMVQNHADWVCVGMFFVL-----GLMFEGETAEMSVFLTLQHG 56
Dy 1 MAFRRR--KSYPLFSQEFVHNHAD-----IGFFVLCVLGLMFEVTAFTAFILPQYN 54
Qy 57 VVPAEGLPSGRSLYHYGVKDLATVFYMLVAIIHATIQEYVLDKLSRRLQTKGKON 116
Dy 55 SSIQT---LGEILYHYGVKDLATVILFYVIAIIHAIHVOEYILDKINKRLSLKVKOS 111
Qy 117 KLNAGQLSVFYIVSGTWGMIIASENCLSDDPTLLKMSQPHNMFTQMFFYISQLAYWF 176
Dy 112 RFNESGGLAAFLHASMFWCLVSVATEGYLSYKPTLWESYHYVLPFOVKFFYLQCLAYWL 171
Qy 177 HSPFLYFQVKRQIDIPGLIYIGLHLFHIGGAYLLYLNHLGLLLMLHYAVELLSSVCS 236
Dy 172 HALPELIFQVKKEVPRLQYIVLYLLHAGAYLLNLRGLILLLLLOSVAEFLFHIAH 231
Qy 237 LLVFGDERYQKGLSLWPVIFSGRLVTLIVSVTVVGLHLAGTN-RNGNALSGNVNVIK 295
Dy 232 LFYETDENNORLFWACWVFEVITRLFTLTLVLTIGFLARAEVHTFNPDKGTLNLLFR 291
Qy 296 IAVLSSCSIOVYITWTTLVTLVWLQWLEDANLHVCGKRRSRKGT----- 342
Dy 292 MVLLLCVQSOTMMWFHIFHQLWRE-----CCKEQAAKRKRSVAVAMKQAKVIKR 345
Qy 343 -----ENGVENPNRIDSPPKKKEKAP 363
Dy 346 ESGYHGVVYKAENGSPRKKIKSP 371

RESULT 5
ID Q92425 ,PRELIMINARY; PRT; 370 AA.
AC Q92425;
DT 01-DEC-2001 (T-EMBLrel. 19, Created)
DT 01-DEC-2001 (T-EMBLrel. 19, Last sequence update)
DE TRAM2.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OC NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RA Hartmann E.;
RL Submitted (APR-2001) to the EMBL/GenBank/DDBJ databases.
DR EMBL; AY029530; AAK40298.1; -.
SQ SEQUENCE 370 AA; 43182 MW; F8E768AFB0582548 CRC64;

Query Match 38.4%; Score 727; DB 11; Length 370;
Best Local Similarity 40.7%; Pred. No. 3.4e-57;
Matches 151; Conservative 70; Mismatches 134; Indels 16; Gaps 5;

Qy 5 KKNARNPPVLSHEFMVQNHADWVCVGMFFVLGLMFEGETAEMSVFLTLQHG VVPAEGL 64
Dy 4 RRRTKSPLESQEFVHNHADIGFCLVLCVLGLMFEVTAFTAFILPQYNISVPT--- 60
Qy 65 PSGSRTL-YHYGVKDLATVFYMLVAIIHATIQEYVLDKLSRRLQTKGKONLEAQ 123
Dy 61 -ADSETVHYHYGPKDLTILFYVITITFIHVAVVOEYILDKISRLSLKVKHKNESQ 119
Qy 124 LSVFYIVSGTWGMIIASENCLSDDPTLLKMSQPHNMFTQMFFYISQLAYWFHSPFLY 183
Dy 120 LLVFLHSAVACVFIIVTEGYLTNPNRSLMEDYPHYVLSYFQVKFFYLGQLAYLHSLPELY 179
Qy 184 FQVKRQIDIPGLIYIGLHLFHIGGAYLLYLNHLGLLLMLHYAVELLSSVCSLLYFGDE 243
Dy 184 FQVKRQIDIPGLIYIGLHLFHIGGAYLLYLNHLGLLLMLHYAVELLSSVCSLLYFGDE 243
```



NA	name:	
NT	genome sequence of the nematode <i>C.elegans</i> : A platform for	
RT	investigating biology.;	
RL	Science 282:2012-2018(1998).	
DR	ENBL; AL022716; CAA18770.1; -.	
DR	InterPro: IPR001185; MSLC.	
SQ	SEQUENCE 371 AA; 42420 MW; 5269FC7231222592 CRC64;	

Qy	2	GLRKNARNPPVLSHEFWQNHADWSCVGMFFVYGLMFECTAEMSVFTLTQGVVVPA	61
	:	:     :     :     :     :     :     :     :     :     :	
Db	7	GSKASKPQPPILSHEPTIQNHGDISCVVMVFVGLMFPPLTHLSLFTAPQNGTYTV	66
Qy	62	EGLPSGSGRTLYHY--GWKDLATVFYMLVAIIHATQIYVLDKLSRRRLQTKGQKNKLN	119



Qy	166	FFYISQLAYWPHSPPELYFOVKRKQDIPGOLIIY----	IGLHFHFIHGGAIIYLNLHLGLLL	221
Dd	181	WYMLEAGFYLSLLILPFD-VKRRDKFKEQVHHFVAVL-	---IGFSYSVNLLRIGAVV	274
Qy	222	LMLHYAVEISSLSSCLLYFGDERYOKGLS-----	LWPIVFTSGRLVTLIYSVVTVGLH	274
Dd	236	LLLHCSDYILEGCIIILNYA--HFRGGCDALFMFALVFYTRLF	FPTQVIYTSVY	290
RESULT 13				
ID	Q9HA82	PRELIMINARY;	PRT;	394 AA.
AC	O9HA82:			
DT	01-MAR-2001	(TrEMBLrel. 16, Created)		
DT	01-MAR-2001	(TrEMBLrel. 16, Last sequence update)		
DE	01-DEC-2001	(TrEMBLrel. 19, Last annotation update)		
DE	HYPOTHETICAL 46.4 KDA PROTEIN.			
OS	Homo sapiens (Human).			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OX	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.			
NCBI_TaxID	9606;			
RN	[1]			
RC	SEQUENCE FROM N.A.			
TISSUE	EMBRYO;			
RA	Isogai T., Ota T., Hayashi K., Sugiyama T., Otsuki T., Suzuki Y.,			
RA	Nishikawa T., Nagai K., Sugano S., Shiratori A., Sudo H., Sugawara M.,			
RA	Takatsuna M., Hosoi R., Kakui Y., Kodaira H., Kondo H., Sugawara M.,			
RA	Tagahashi M., Chiba Y., Ishida S., Murakawa K., Ono Y., Takiguchi S.,			
RA	Watanabe S., Kimura K., Murakami K., Ishii S., Kawai Y., Saito K.,			
RA	Yamanoto J., Wakamatsu A., Nakamura Y., Nagahari K., Masuno Y.,			
RA	Ninomiya K., Iwayanagi T.;			
RT	*NEBO human cDNA sequencing project.*;			
RL	Submitted (AUG-2000) to the EMBL/GenBank/DDBJ databases.			
RZ	[2]			
RN	SEQUENCE FROM N.A.			
RC	TISSUE=LUNG CARCINOMA;			
RA	Strausberg P.;			
RA	Submitted (JUL-2001) to the EMBL/GenBank/DDBJ databases.			
DR	EMBL; AK022151; BAB19972.1; -			
DR	EMBL; BC009828; AAA09828.1; -			
DR	InterPro: IPRO01356; Homeobox.			
DR	SMART; SM00389; HOX; 1.			
DR	PROSITE; PS50071; HOMEBOX_2; 1.			
KW	Hypothetical protein.			
SEQ	SEQUENCE 394 AA; 46398 MW; 67228829BDED2801 CRC64;			
Query Match 7.3%; Score 138.5; DB 4; Length 394;				
Best Local Similarity 24.7%; Pred. No. 0.00027;				
Matches 44; Conservative 42; Mismatches 81; Indels 11; Gaps				
Qy	107	RLQTKGKNKLNEAGQLSVFYGIWGMIILASENCLSDPTLLWKKSOPHNMTFOKF	166	
Dd	126	RPOLTK----KFCESWRFLFYLUSSVGSLVGHESWLAPVMCMWDVPNTLKPSLYW	181	
Qy	167	FYISQLAYWPHSPPELYFOVKRKQDIPGOLIYIGLHFHFIHGGAIIYLNLHLGLLLMLHY	226	
Dd	182	WYMLEAGFYLSLLILPFD-VKRRDKFKEQVHHFVAIVLTFTSYSANLLRIGSLVLLHD	240	
Qy	227	AVELSSVCCLLYFGDERYOKGLS-----	LWPIVFTSGRLVTLIYSVVTVGLHCACTNR	280
Dd	241	SSDIYLEACKMNVY--MQQQQCVDALFLIFSFFVYTRLVLPFTQILTYYTESISR	296	
RESULT 14				
ID	Q9M6A4	PRELIMINARY;	PRT;	308 AA.
AC	Q9M6A4:			
DT	01-OCT-2000	(TrEMBLrel. 15, Created)		
DT	01-OCT-2000	(TrEMBLrel. 15, Last sequence update)		
DE	01-OCT-2000	(TrEMBLrel. 15, Last annotation update)		
DE	ASCI.			

GN AS.

OS Lycopersicon esculentum (Tomato).

OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;

OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;

OC Asterales; euasterids I; Solanales; Solanaceae; Solanum.

OX NCBI\_TaxID=4081,

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=cv. VFNT CHERRY;

RX MEDLINE=99168767; PubMed=10071209;

RA Mesbah L.A., Kneppers T.J., Takken F.L., Laurent P., Hille J.,

RA Nikkamp H.J.;

RT "Genetic and physical analysis of a YAC contig spanning the fungal

RT disease resistance locus Asc of tomato (*Lycopersicon esculentum*).";

RL Mol. Gen. Genet. 261:50-57(1999).

RN [2]

RP SEQUENCE FROM N.A.

RC STRAIN=cv. VFNT CHERRY;

RX MEDLINE=20243803; PubMed=107811105;

RA Brandwagt B.F., Mesbah L.A., Takken F.L.W., Laurent P.L.,

RA Kneppers T.J.A., Hille J., Nikkamp H.J.;

RT "A longevity assurance gene homolog of tomato mediates resistance to

RT Alternaria alternata f. sp. lycopersici toxins and fumonisins B1.";

RT Proc. Natl. Acad. Sci. U.S.A. 97:4961-4966(2000).

DR ENBL; AF198177; AAF67518.1;

SQ SEQUENCE 308 AA; 36301 MW; 2BD3378CE53F416D CRC64;

Query Match 6.9%; Score 131; DB 10; Length 308;  
Best Local Similarity 19.5%; Pred. NO. 0.00094;  
Matches 46; Conservative 50; Mismatches 80; Indels 60; Gaps

QY 62 EGLPGSGSRTLYHYGVKDLAVFYFMVLVAIIHATIQEYLDKLRRLLQLTK-----L112  
| | | : | : | : | : | : | : | : | : | : | : | : | :  
Db 16 ESUPE-----YODLIFFLEFPVILRFDLFVEALAKRMIFGKTVVNINGR 66

QY 113 ---GQKNLNAGOLSVFYIVSWGIMGMILAENCLSDPTLLWKSO----PHNMFTQM 164  
| | | : | : | : | : | : | : | : | : | : | : | : | :  
Db 67 EERKKINKFKESAWKFYFIETLSAEALLSVFCNEPWFSTDSEFYWAGPDVWPMLKMKL 136

QY 165 KFFVISOLAWEHS-PPELVFOVKRDIPGLIYIGLHLFGGAYLLNH-----L 212  
| : | : | : | : | : | : | : | : | : | : | : | : | : | :  
Db 127 KLLYNWAGGYFFYSIFATLWE--TRRYDFAAQH-----HVTVSLIVLSVYGARI 179

QY 218 GLLLLMLHYAVEILLSVCSL-----LYFGDERYOGKLSMPVFISGRLV 262  
| : | : | : | : | : | : | : | : | : | : | : | : | : | :  
Db 180 GSVVALHDGDGVMEIAKMSKYSGFDLIAIFF-----SLFAVTSLRII 226

RESULT 15

ID Q95RN6 PRELIMINARY; PRT; 400 AA.

AC Q95RN6;

DT 01-DSC-2001 (TrEMBLrel. 19, Created)

DD 01-DSC-2001 (TrEMBLrel. 19, Last sequence update)

DE D18904P.

GS CG15898.

OS Drosophila melanogaster (Fruit fly).

OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;

OC Pserygota; Neoptera; Diptera; Brachycera; Muscomorpha;

OC Ephydroidea; Drosophilidae; Drosophila.

OX NCBI\_TaxID=7227;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=y, CN BW SP;

RA Stapleton M., Brokstein P., Hong L., Agbayani A., Carlson J.,

RA Champe M., Chavez C., Dorsett V., Farfan D., Frise E., George R.,

RA Gonzalez M., Guarin H., Li P., Liao G., Miranda A., Mungall C.J.,

RA Nunoo J., Paclzel J., Paragas V., Park S., Phouanenavong S., Wan K.,

RA Yu C., Lewis S.E., Rubin G.M., Ceinaliker S.;

RL Submitted (OCT-2001) to the EMBL/GenBank/DDJB databases.  
DR ENBL; AY061255; AAL28803.1; -.

Search completed: September 20, 2002, 14:38:59  
Job time: 246 sec